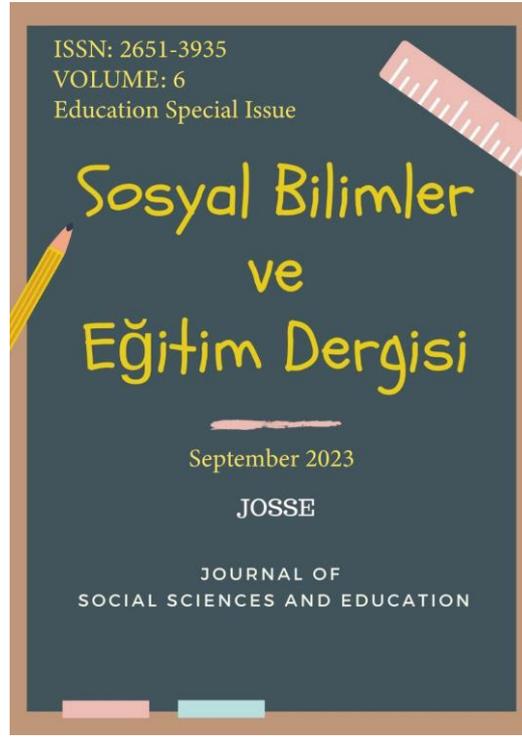


JOURNAL OF SOCIAL SCIENCES AND EDUCATION (JOSSE)



<https://dergipark.org.tr/tr/pub/josse>

Analytical Rubric Development Study for A2 Level Speaking Skills in Teaching Turkish as a Foreign Language

Eda TEKİN¹

Yildiz Technical University, Faculty of Education, Teaching Turkish Language Department

Research Assist. Dr.

edatkn@hotmail.com

Orcid ID: 0000-0001-6831-8101

Muhammed Eyyüp SALLABAŞ²

Yildiz Technical University, Faculty of Education, Teaching Turkish Language Department

Prof. Dr.

sallabas@gmail.com

Orcid ID: 0000-0003-4346-4385

Article Type: Research Article

Received: 20.08.2023

Revision received: 22.09.2023

Accepted: 22.09.2023

Published online: 25.09.2023

Citation: Tekin, E. & Sallabaş, M. E. (2023). Analytical rubric development study for A2 level speaking skills in teaching Turkish as a foreign language. *Journal of Social Sciences and Education*, 6(Education Special Issue), 362-379.

Analytical Rubric Development Study for A2 Level Speaking Skills in Teaching Turkish as a Foreign Language

Eda TEKİN¹

Yildiz Technical University, Faculty of Education, Teaching Turkish Language Department

Muhammed Eyyüp SALLABAŞ²

Yildiz Technical University, Faculty of Education, Teaching Turkish Language Department

ABSTRACT

Research Article

Speaking skill is one of the most important skills in learning a foreign language and realizing the act of communication. The more the student improves in speaking skills, the more self-confidence will increase. In the development of speaking skill, measurement and evaluation of the skill is as important as all the activities carried out during the lesson. It is important for the students to receive feedback on their speaking skill and to know how scoring is done. In this research, it is aimed to develop a rubric that will guide both students and teachers in order to evaluate the speaking performance of students at the A2 level of teaching Turkish as a foreign language. In the study, the validity and reliability studies of the rubric were carried out. A group of four field experts was formed in the study. As a result of the scores made by four experts of the field, the reliability between the raters was calculated according to the Intraclass Correlation statistics in the SPSS 25 Program. According to the feedback from the experts, it has been revealed that it is a tool that both teachers and students can use as an assessment tool. It is thought that the rubric will contribute to the researchers while creating a tool at the same level or at other levels, and will guide the teachers and learners in the stages of measuring and evaluating the A2 level speaking skill of teaching Turkish as a foreign language.

Received: 20.08.2023

*Revision received:
22.09.2023*

Accepted: 22.09.2023

*Published online:
25.09.2023*

Keywords: Turkish as a foreign language, A2 level, speaking skill, rubric, validity and reliability

¹ Corresponding author:

Research Assist. Dr.

edatkn@hotmail.com

Orcid ID: 0000-0001-6831-8101

² *Prof. Dr.*

sallabas@gmil.com

Orcid ID: 0000-0003-4346-4385

Introduction

Speaking skill, which is the most important skill for the realization of communication in foreign language teaching, is one of the four basic skills. Reading and listening skills help the person understand themselves and their surroundings, while speaking skills are necessary to express thoughts verbally using language and initiate the communication process (Yorgancı & Baş, 2021, pp. 70). Language skills are described in the Common European Framework of Reference for Languages CEFR (2021) and are used to identify, develop and update the competencies of foreign language teaching for every skill and level.

The most basic principle of foreign language teaching is to teach language for communication. Therefore, the ability to speak among the four basic language skills is the most important (İşisağ & Demirel, 2010, pp. 193). In language education, it is important to give students the skills to understand and explain the target language. The ability to speak is the most widely used skill in everyday life apart from other skills. Therefore, the main goal of learning speech skills in language education is to enable students to express emotions and thoughts fluently using the language they learn (Koçak, 2018, pp. 17). The importance of speaking in any foreign language is important, and evaluating the ability to speak can also be problematic. Measurement and evaluation occupy an important place in the educational process in order for the student to learn about his or her situation, to be able to develop himself, and continue the process as they develop. Field in the assessment and evaluation of language skills in foreign language teaching are an area of research that attracts attention and a need. Speaking skill, in particular, stand out as one of the least studied subjects in Turkish teaching as a foreign language, so it seems that more research on these skills needs to be done in the studies to be conducted (Arıcı et al., 2017). It is of utmost importance to use the analytical section to correctly measure speaking skills in Turkish as a foreign language teaching, so as not to differ from variables such as teachers, institutions studied and used textbooks (Boylu, 2019, pp. 144).

Today, along with the constructivist approach, the dimensions of measurement and evaluation have also changed, replacing traditional techniques with more modern and process-based measuring and assessment techniques (Yılmaz, 2018, pp. 1625) and one of them is rubric. It is divided into two categories, holistic and analytical. (Chase, 1999; Mertler, 2001; Nitko, 2001). Since the assessment of overall performance is key, the overall rating keys are used when the purpose of the performance assessment is summary. As a result of this ranking

of performance tasks with holistic rubrics, the student is given only limited feedback (Mertler, 2000, pp. 1). The analytical rubrics are more comprehensive and important in terms of giving students feedback. The use of analytical rubrics represent assessment at a multi-dimensional level (Mertler, 2001).

Teachers are in a very critical position as those who organize course objectives in educational teaching environments, motivate students and evaluate the measurement with course outcomes (İnal, 2020, pp. 191). The absence of a common rating tool in the assessment of speech skills in the field of teaching Turkish as a foreign language is a missing issue. Failure to use a common analytical rubric among Turkish educational centers can cause reliability and validity problems. In today's assessment practices, a valid analytical rubric is not used. Some institutions use a single evaluator in speech exams, while others use their own rubrics (Kahveci, 2022, pp. 13).

Speaking skill is one of the skills that students have the most difficulty with in foreign language teaching, and how this skill is measured and assessment and evaluation is an important issue for both teachers and learners. Rubrics serve as a guide so that students' performance can be measured more objectively and clearly. It can be said that analytical rubrics are more reliable, especially since they are more detailed and descriptive of each item (Mertler, 2001). In the development of narrative skills (speaking and writing), it is expected that the skills will be realized as practice. As a result of this application, a performance emerges and rubrics (analytic scales, holistic scales and observation forms) are frequently used in performance-based evaluations (Bozkurt & Arıca-Akkök, 2019, pp. 419). In addition, rubrics are a necessity when an adult language learner wants to get comprehensive and detailed feedback on his speaking skill. Rubrics have an important place in the literature and practice in order to not only score the students' speeches, but also to know according to which criteria they are evaluated and to explain a more objective scoring to the students. Assessment and evaluation is an important subject in every field of education and should be evaluated according to every skill level and criteria in foreign language teaching. As Boylu (2019) stated, the arbitrary (based on experience) evaluation of speaking and writing skills by teachers is one of the assessment and evaluation problems.

Rubrics guide us towards our goals as teachers. We use them to clarify our learning goals, design guidelines that address those goals, communicate goals to students, guide our feedback on students' progress toward goals, and judge end products by the degree to which

goals are achieved. Like many teachers, I use the instructions before, during and after teaching and it has many benefits (Andrade, 2005, pp. 27).

Validity, reliability and objectivity are important issues for rubrics. To be valid, a teaching rubric must, at a minimum, comply with reasonable and respected standards and the curriculum being taught. When used by different people, it should pass the reliability test by giving similar ratings (Andrade, 2005, pp. 30). With this research, it is aimed to develop an analytical rubric that can be used in the evaluation of A2 level speaking skills, independent speaking and conversational speaking skills for educators and students in the field of Turkish as a foreign language. In addition, it is thought that a rubric for the A2 level of speaking skill will be developed and its validity and reliability will be revealed, and it will guide future research and its development at other levels.

Method

Research Participants

The participants of the study consisted of Turkish as a foreign language teachers working at the Turkish Teaching Application and Research Center within a state university and foreign students learning Turkish at the A2 level of the same university. The research was carried out with a study group consisting of four field expert lecturers and twenty-four A2 level students.

Data Collection Tools

In this study, A2 Level Speaking Proficiency Exam was created in advance to receive student answers orally. The final versions of the exam questions were given by taking the opinions of four field experts. In addition, A2 Level Speaking Analytical Rubric for Turkish Learners as a Foreign Language was used as a data collection tool to evaluate students' A2 level speaking skills. This rubric has been prepared by reviewing the literature, taking into account the criteria specified in the Common European Framework of Reference for Languages (2021). The rubric is defined separately for independent and conversational skills. The rubric was created in its final form by referring to expert opinions, and validity and reliability analyzes were made.

Research Context

In this study, it is aimed to prepare an analytical rubric to evaluate the speaking skills of students at A2 level of teaching Turkish as a foreign language. The criteria to be included in the rubric were primarily determined according to the A2 level speaking skills of the Common European Framework of Reference for Languages (2021), and were defined separately for independent speaking and conversational skills. Then, the rubric was finalized by taking expert opinions on the rubric items.

The data of the study were obtained by evaluating the video-recorded speeches of twenty-four students by four experts of the field. Rubric items were determined on the basis of literature review, competencies in the Common European Framework of Reference for Languages, and expert opinions. Finally, the content validity index was calculated by submitting the draft rubric to the opinions of field experts.

Scores from 1 to 3 are included in the scale. The created scale was tested on the recorded speeches of A2 level students studying at Yıldız Technical University TOMER. The aim of the research is to develop a rubric for the assessment of A2 level speaking skills for learners of Turkish as a Foreign Language. This rubric was first developed by scanning the relevant literature. Later, CEFR (2021) A2 level competencies were determined. As rubric preparation criteria, the studies of various researchers such as Andrade (1997), Popham (1997), Moskal and Leydens (2000), Mertler (2001), Andrade (2005) were examined. These criteria were combined and rubric preparation steps were applied.

1. Relevant literature review
2. Creating the items to be included in the rubrics
3. Preparation of two separate draft rubrics (Independent speech and conversational)
4. Sending rubrics to expert opinion
5. Use of draft rubrics
6. Getting feedback from experts
7. Editing rubrics
8. Conducting reliability and validity studies of rubrics
9. Finalizing the rubrics

Since the purpose of this research was to create an analytical rubric to evaluate students' speaking skills, A2 level speaking competency criteria of the Common European Framework of Reference for Languages (2021) were used. As a result of the literature review, rubric items were defined separately for both independent and conversational skills.

Afterwards, a rubric was developed by taking expert opinions. After expert opinions, the rubric was finalized and its validity and reliability were examined.

Analysis of Data

The validity of the rubric created in the research was provided in line with the opinions of four experts on the subject. Validity relates to how accurately the measure measures the feature it is particularly interested in, by distinguishing it from other features. In other words, the validity of the measurement results refers to the degree to which the measurement is aimed accurately (Büyükozturk et al., 2020, pp. 111).

In practice, students were asked four questions for independent speaking and they were asked to speak. Five questions were asked to the students for conversational speaking and they were asked to speak. The conversations were recorded with audio and video. In the applications, four experts listened to each participant and evaluated the speeches. Intraclass Correlation coefficient of agreement was used for inter-rater reliability. The concept of reliability generally refers to the level of consistency between analyzes of data sets by different coders in qualitative research (Creswell, 2021).

Compliance with Ethical Standard

The study was approved by the Social and Human Sciences Research Ethics Committee of Yildiz Technical University (Date: 26.08.2022, Session No: 2022.08).

Findings

Findings Regarding the Validity of the Developed Analytical Rubric

In the qualitative research carried out with the participation of four experts, the analytical rubric, which was prepared to measure the speaking skills of Turkish learners as a foreign language, was evaluated by experts and examined in terms of its suitability. Using the Lawshe analysis method (Yurdugül, 2005), the experts were asked to evaluate the items and their contents in the rubric according to triple criteria such as "appropriate", "partially appropriate" and "not suitable/explanation". According to expert opinions, the content validity index of the items is explained in Table 1.

Table 1

Content Validity Index for Independent Speech Items

| A2 Level Speaking Skill Independent Speaking Items | A | PA | NS | CVI |
|--|------|----|----|------|
| Item One (Introduction) | 4 | 0 | 0 | 1.00 |
| Item Two (Main Idea) | 4 | 0 | 0 | 1.00 |
| Item Three (Conclusion) | 4 | 0 | 0 | 1.00 |
| Item Four (Vocabulary) | 4 | 0 | 0 | 1.00 |
| Item Five (Grammar Mistakes) | 4 | 0 | 0 | 1.00 |
| Item Six (Fluency) | 4 | 0 | 0 | 1.00 |
| Item Seven (Pronunciation) | 4 | 0 | 0 | 1.00 |
| Item Eight (Stress and Intonation) | 4 | 0 | 0 | 1.00 |
| Item Nine (Gesture and Mimics) | 4 | 0 | 0 | 1.00 |
| Item Ten (Speaking Speed) | 4 | 0 | 0 | 1.00 |
| Number of Experts | 4 | | | |
| Content Validity Index (CVI) | 1.00 | | | |

A = Appropriate, PA = Partially Appropriate, CVI= Content Validity Index

Four field experts state that the items in question are appropriate in terms of content validity. This shows that the measurement tool is reliable in terms of content validity and will be effective in measuring the skills it aims at.

Table 2

Content Validity Index for Conversation Items

| A2 Level Conversation Skills Conversation Items | A | PA | NS | CVI |
|---|---|----|----|------|
| Item One (Introduction) | 4 | 0 | 0 | 1.00 |
| Item Two (Main Idea) | 4 | 0 | 0 | 1.00 |
| Item Three (Conclusion) | 4 | 0 | 0 | 1.00 |
| Item Four (During Speaking) | 4 | 0 | 0 | 1.00 |
| Item Five (Vocabulary) | 4 | 0 | 0 | 1.00 |

| | | | | |
|---|------|---|---|------|
| Item Six (Grammar Mistakes) | 4 | 0 | 0 | 1.00 |
| Item Seven (Fluency) | 4 | 0 | 0 | 1.00 |
| Item Eight (Pronunciation) | 4 | 0 | 0 | 1.00 |
| Item Nine (Stress and Intonation) | 4 | 0 | 0 | 1.00 |
| Item Ten (Gesture and Mimics) | 4 | 0 | 0 | 1.00 |
| Item Eleven (Speaking Speed) | 4 | 0 | 0 | 1.00 |
| Item Twelve (Maintaining Communication) | 4 | 0 | 0 | 1.00 |
| <hr/> | | | | |
| Number of Experts | 4 | | | |
| Content Validity Index (CVI) | 1.00 | | | |

A = Appropriate, PA = Partially Appropriate, CVI= Content Validity Index

Four field experts state that the items in question are appropriate in terms of content validity. This shows that the measurement tool is reliable in terms of content validity and will be effective in measuring the skills it aims at.

Findings Regarding the Reliability of the Developed Analytical Rubric

Consistency of evaluation scores is an important factor affecting the reliability of measurement. In a reliable test, a student expects the same result regardless of the time their answers are scored and the evaluator. However, in an unreliable exam, the student's score may change depending on factors other than the objectives of the exam (Moskal & Leydens, 2000, pp. 4). During the rubric development process, the exams of 24 participants were evaluated separately by four different field experts, and the inter-rater consistency coefficients were calculated for each item for the data obtained. The consistency of the assessments made by the raters on the student papers for each item was tested with the Intraclass Correlation analysis. The Intraclass Correlation coefficient is used to measure the repeatability of a measuring instrument and whether measurements made by different gauges or at different times give the same results. Therefore, high values of the Intraclass Correlation coefficient indicate that the measuring instrument is reliable. The Intraclass Correlation coefficient takes values between 0 and 1. Values less than 0.5 indicate poor reliability, values between 0.5 and 0.75 indicate moderate reliability, values between 0.75 and 0.9 indicate good reliability, and values greater than 0.90 indicate excellent reliability (Koo & Li, 2016). The obtained

consistency coefficients show the reliability of the relevant rubric. The inter-rater reliability coefficient for each item is presented in Table 3 and Table 4.

Table 3

Inter-Rater Reliability Intraclass Correlation Coefficients of Independent Speech

| Items | Measurement | Inter-Rater Correlation | %95 Confidence Interval | |
|-------|---------------------|-------------------------|-------------------------|-------------|
| | | | Lower Bound | Upper Bound |
| 1 | Single Measurements | .940 ^a | .893 | .971 |
| | Mean Measurements | .984 ^c | .971 | .993 |
| 2 | Single Measurements | .923 ^a | .863 | .962 |
| | Mean Measurements | .979 ^c | .962 | .990 |
| 3 | Single Measurements | .896 ^a | .819 | .948 |
| | Mean Measurements | .972 ^c | .948 | .987 |
| 4 | Single Measurements | .884 ^a | .799 | .942 |
| | Mean Measurements | .968 ^c | .941 | .985 |
| 5 | Single Measurements | .926 ^a | .868 | .963 |
| | Mean Measurements | .980 ^c | .963 | .991 |
| 6 | Single Measurements | .916 ^a | .868 | .964 |
| | Mean Measurements | .980 ^c | .964 | .991 |
| 7 | Single Measurements | .853 ^a | .750 | .925 |
| | Mean Measurements | .959 ^c | .923 | .980 |
| 8 | Single Measurements | .871 ^a | .779 | .935 |
| | Mean Measurements | .964 ^c | .934 | .983 |
| 9 | Single Measurements | .921 ^a | .860 | .961 |
| | Mean Measurements | .979 ^c | .961 | .990 |
| 10 | Single Measurements | .908 ^a | .838 | .954 |
| | Mean Measurements | .975 ^c | .954 | .988 |

As a result of the reliability analysis, it was determined that the reliability level of all items was in the category of "excellent".

Table 4

Inter-rater Reliability Intraclass Correlation Coefficients of Conversational Speaking

| Items | Measurement | Inter-Rater Correlation | %95 Confidence Interval | |
|-------|---------------------|-------------------------|-------------------------|-------------|
| | | | Lower Bound | Upper Bound |
| 1 | Single Measurements | .940 ^a | .893 | .971 |
| | Mean Measurements | .984 ^c | .971 | .993 |
| 2 | Single Measurements | .923 ^a | .863 | .962 |
| | Mean Measurements | .979 ^c | .962 | .990 |
| 3 | Single Measurements | .896 ^a | .819 | .948 |
| | Mean Measurements | .972 ^c | .948 | .987 |
| 4 | Single Measurements | .924 ^a | .865 | .963 |
| | Mean Measurements | .980 ^c | .962 | .990 |
| 5 | Single Measurements | .884 ^a | .799 | .941 |
| | Mean Measurements | .968 ^c | .942 | .985 |
| 6 | Single Measurements | .926 ^a | .868 | .963 |
| | Mean Measurements | .980 ^c | .963 | .991 |
| 7 | Single Measurements | .926 ^a | .868 | .964 |
| | Mean Measurements | .980 ^c | .964 | .991 |
| 8 | Single Measurements | .853 ^a | .750 | .925 |
| | Mean Measurements | .959 ^c | .923 | .980 |
| 9 | Single Measurements | .871 ^a | .779 | .935 |
| | Mean Measurements | .964 ^c | .934 | .983 |
| 10 | Single Measurements | .921 ^a | .860 | .961 |
| | Mean Measurements | .979 ^c | .961 | .990 |
| 11 | Single Measurements | .908 ^a | .838 | .954 |

| | | | | |
|----|---------------------|-------------------|------|------|
| | Mean Measurements | .975 ^c | .954 | .988 |
| 12 | Single Measurements | .917 ^a | .853 | .959 |
| | Mean Measurements | .978 ^c | .959 | .989 |

As a result of the reliability analysis, it was determined that the reliability level of all items was in the category of "excellent".

Discussion and Results

Analytical rubrics in foreign language teaching are tools that help teachers to objectively evaluate their students' written and oral performances and to improve students' language skills by giving feedback. Analytical rubrics allow students to analyze in detail their performance on specific language skills (eg grammar, vocabulary, expression, pronunciation, etc.). This tool helps teachers identify their students' strengths and weaknesses and improve their skills by giving students appropriate feedback (Ulker, 2017; Vercellotti & McCormick, 2021). Analytical rubrics also help teachers analyze and grade students' written and oral performances in detail.

In this study, an analytical rubric was developed to evaluate the speaking skills of A2 level students learning Turkish as a foreign language. The validity and reliability of the developed rubrics were also examined within the scope of the research. The developed rubric includes ten criteria in independent speaking to measure the speaking skills of A2 level students; It includes twelve criteria in conversation. It should be noted that the rubric is a rubric that can be used for A2 level, since the developed rubric criteria are created according to the CEFR (2021) A2 level and the competencies in the books used in teaching Turkish as a foreign language. The analytical rubric was used to evaluate students' speaking skills in detail. The rubric provides an objective measurement of students' speaking skills by scoring the criteria under different categories separately. The developed analytical rubric has been tested with validity and reliability analyzes. According to the results of the validity analysis, the rubric accurately measures the speaking skills of the students. In addition, the results of the reliability analysis show that the rubric has a high reliability in scoring. These findings show

that analytical rubric is an effective tool for objective and reliable evaluation of speaking skills of students learning Turkish as a foreign language.

With the developed rubric, teachers can benefit from the evaluation of students' performance and exams, and students can develop an awareness of speaking skill processes. In order to evaluate students themselves, educational rubrics should be written in a language that students can understand, that is, they should be understandable, describe the quality of the work, include general weaknesses and how to avoid them, and be used as a guide by students (Andrade, 2001, pp. 1). Analytical rubrics help teachers teach and evaluate specific skills to enable students to achieve their learning goals. As a result, analytical rubrics are an effective tool in foreign language teaching that helps teachers to objectively evaluate their students' performances and improve their skills by giving feedback.

Assessment and evaluation is also an important issue in teaching Turkish as a foreign language and various studies have been conducted (Boylu, 2019; Gedik, 2017; Karagöl, 2020; Kesici, 2022), but assessment tools for skills and levels are limited. In this context, it is important to create separate rubrics for each skill and for each level in order to make measurement and evaluation more objective. It is noteworthy that more practice will be beneficial in teaching Turkish as a foreign language, especially in evaluating and developing speaking skills, and it is thought that the developed analytical rubric will contribute to future studies.

Recommendations

- The developed analytical rubric can contribute to both teachers and students in assessing speaking skills more objectively.
- The rubric can serve as a guide for other Turkish language teachers to conduct effective assessments.
- The rubric can be employed to track students' progress in speaking skills. Teachers can enhance student motivation by providing regular feedback and clearly communicating goals.
- The rubric opens doors for new research in the field of Turkish language teaching such as exploring speaking skills measurement and assessment further by using the rubric as a foundation.

- The rubric can be applied internationally for assessing Turkish language proficiency.

Compliance with Ethical Standard

The study was approved by the Social and Human Sciences Research Ethics Committee of Yildiz Technical University (Date: 26.08.2022, Session No: 2022.08).

References

- Andrade, H. G. (1997). Understanding rubrics. *Educational leadership*, 54(4), 14-17.
- Andrade, H. G. (2001). The effects of instructional rubrics on learning to write. *Current issues in education*, 4. <https://cie.asu.edu/ojs/index.php/cieatasu/article/view/1630>
- Andrade, H. G. (2005). Teaching with rubrics: The good, the bad, and the ugly. *College teaching*, 53(1), 27-31. <https://doi.org/10.3200/CTCH.53.1.27-31>
- Arıcı, A. F., Sallabaş, M. E., & Başaran, M. (2017). Speaking skills of the learners who learned turkish as a foreign language at B2 Level: A Study on developing the graded scoring scale. *International Online Journal of Educational Sciences*, 9(2), 519-526. <https://doi.org/10.15345/iojes.2017.02.016>
- Avrupa Konseyi. (2021). Diller için Avrupa ortak öneriler çerçevesi: Öğrenim, öğretim ve değerlendirme. MEB Yayınları. http://ttkb.meb.gov.tr/meb_iys_dosyalar/2022_01/04144518_CEFR_TR.pdf
- Boylu, E. (2019). *Yabancılara Türkçe öğretiminde ölçme değerlendirme uygulamaları ve standart oluşturma* [Doktora tezi]. Çanakkale Onsekiz Mart Üniversitesi.
- Bozkurt, B. Ü., & Akkök, E. A. (2019). Anadili Türkçe olan konuşucular için konuşma becerisi derecelendirme ölçeğinin geliştirilmesi. *İlköğretim Online*, 18(1), 416-436. <https://doi.org/10.17051/ilkonline.2019.527649>
- Büyüköztürk, Ş., Kılıç-Çakmak, E., Akgün, Ö., Karadeniz, Ş., & Demirel, F. (2020). *Bilimsel araştırma yöntemleri*. Pegem Akademi.
- CEFR (2021). Diller için avrupa ortak başvuru metni: Öğrenme, öğretme ve değerlendirme. http://ttkb.meb.gov.tr/meb_iys_dosyalar/2022_01/04144518_CEFR_TR.pdf
- Chase, C. I. (1999). *Contemporary assessment for educators*. Longman.

- Creswell, J. W. (2021). *Nitel araştırma yöntemleri*. Qualitative research methods]. M. Bütün & SB Demir, Trans. (Eds.). Siyasal Kitapevi.
- Gedik, E. (2017). *Yabancı dil olarak Türkçe öğretiminde ölçme ve değerlendirme* [Yüksek lisans tezi]. İstanbul Arel Üniversitesi.
- İnal, E. (2020). Yabancı dil öğretiminde konuşma becerisinin geliştirilmesi ve ölçme değerlendirme süreçlerine yönelik düşünceler: öğretici yeterlikleri. *Aydın Tömer Dil Dergisi*, 5 (2), 189-204. <https://dergipark.org.tr/en/pub/aydintdd/issue/60730/851465>
- İşisağ, K. U., & Demirel, Ö. (2010). Diller için Avrupa Ortak Başvuru Metni'nin konuşma becerisinin gelişiminde kullanılması. *Eğitim ve Bilim*, 35 (156).
- Kahveci, N. (2022). *Yabancı dil olarak Türkçe eğitiminde yazma becerisini değerlendirmeye yönelik dereceli puanlama anahtarı geliştirme çalışması* [Yüksek lisans tezi]. Bartın Üniversitesi.
- Karagöl, E. (2020). Proficiency exams in teaching Turkish as a foreign language in TÖMER (Turkish and foreign languages research and application centers). *Journal of Language and Linguistic Studies*, 16(2), 930-947. <https://doi.org/10.17263/jlls.759347>
- Kesici, S. (2022). Yabancılarla Türkçe öğretiminde ölçme ve değerlendirme: öğretmen algıları üzerine bir inceleme. *Uluslararası Filoloji Bengü*, 2 (2), 45-69. 10.29228/filolojibengu.21
- Koçak, S. (2018). *Yabancı dil olarak Türkçe öğretiminde temel düzeyde (A1-A2) konuşma izleni örneğinin hazırlanması*. [Yüksek lisans tezi]. Hacettepe Üniversitesi.
- Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, 15(2), 155-163. <https://doi.org/10.1016/j.jcm.2016.02.012>
- Mertler, C. A. (2000). Designing scoring rubrics for your classroom. *Practical Assessment, Research, and Evaluation*, 7(1), 25. <https://doi.org/10.7275/gcy8-0w24>
- Mertler, C. A. (2001). Using performance assessment in your classroom [Unpublished manuscript]. Bowling Green State University.
- Moskal, B. M., & Leydens, J. A. (2000). Scoring rubric development: Validity and reliability. *Practical Assessment, Research, and Evaluation*, 7(1), 10. <https://doi.org/10.7275/q7rm-gg74>
- Nitko, A. J. (2001). Educational assessment of students. Upper Saddle River, NJ: Merrill.
- Popham, J. W. (1997). What's wrong-and what's right-with rubrics. *Educational Leadership*, 55(2), 72-75.

- Ulker, V. (2017). The design and use of speaking assessment rubrics. *Journal of Education and Practice*, 8(32), 135-141.
- Vercellotti, M. L., & McCormick, D. E. (2021). Constructing analytic rubrics for assessing open-ended tasks in the language classroom. *TESL-EJ*, 24(4), <https://files.eric.ed.gov/fulltext/EJ1288720.pdf>
- Yılmaz, İ. (2018). Yabancılara Türkçe deyimlerin öğretiminde tamamlayıcı ölçme ve değerlendirme yöntemlerinin kullanılması. *Kastamonu Eğitim Dergisi*, 26 (5), 1623-1641. <https://doi.org/10.24106/kefdergi.389782>
- Yorgancı, O. K., & Baş, B. (2021). Yabancı dil olarak türkçe öğrenenler için b1 düzeyi yazma becerisi dereceli puanlama anahtarı geliştirme çalışması. *Okuma Yazma Eğitimi Araştırmaları*, 9(1), 67-80. <https://doi.org/10.35233/oyea.934684>
- Yurdugül, H. (2005). Ölçek geliştirme çalışmalarında kapsam geçerliği için kapsam geçerlik indekslerinin kullanılması. XIV. Ulusal Eğitim Bilimleri Kongresi, 1, 771-774.

Additional 1

Yabancı Dil Olarak Türkçe Öğrenenler için A2 Düzeyi Bağımsız Konuşma Becerisi Dereceli Puanlama Anahtarı

| | Maddeler | 1 | 2 | 3 |
|-----------------------------------|-------------------------------|--|--|---|
| Konuşmanın İçeriği | Başlangıç | Konuşmaya uygun ifadelerle başlamamıştır. | Konuşmaya kısmen uygun ifadelerle başlamıştır. | Konuşmaya uygun ifadelerle başlamıştır. |
| | Ana Düşünce | Konuşmasında ana düşünceye yer vermemiştir. | Konuşmasında ana düşünceye kısmen yer vermiştir. | Konuşmasında ana düşünceye yer vermiştir. |
| | Bitiş | Konuşmayı sonlandıracak ifadeleri kullanmamıştır. | Konuşmayı sonlandıracak ifadeleri kısmen kullanmıştır. | Konuşmayı sonlandıracak ifadeleri kullanmıştır. |
| Söz Varlığı | Kelime Hazinesi Durumu | Kelime hazinesi yetersizdir. | Kelime hazinesi sınırlıdır. | Kelime hazinesi yeterlidir. |
| Konuşmanın Şekilsel boyutu | Dilbilgisel Hatalar | Konuşurken hatalı dilbilgisel yapılar kullanmıştır (Yaptığı konuşmasının anlaşılmasını etkilemektedir.) | Konuşurken hatalı dilbilgisel yapılar kullanması konuşmasının anlaşılmasını kısmen etkilemektedir. (Konuşurken dil bilgisel hataları vardır ancak ne söylemek istediği anlaşılmaktadır.) | Konuşurken hatalı dilbilgisel yapılar kullanımı yok denecek kadar azdır. (Konuşurken anlaşılmasını etkileyecek dil bilgisel hatalar yapmamıştır.) |
| | Akıcılık | Konuşması akıcı değildir. Çok fazla gereksiz duraklamalar ve kelime seçimlerinde tereddütler yaşamıştır. | Konuşması kısmen akıcıdır. Kısmen daha az duraklama ve tereddüt yaşamıştır. | Konuşması akıcıdır. Konuşurken gereksiz duraklamalar yapmamıştır. |
| | Telaffuz | Telaffuz hataları çok fazladır. Telaffuzu anlaşılır değildir. | Telaffuz hataları göreceli daha azdır. Telaffuzu kısmen anlaşılabilir. | Telaffuz hataları yok denecek kadar azdır. Telaffuzu anlaşılır ölçüdedir. |
| | Vurgu ve Tonlama | Konuşma esnasında vurgu ve tonlamaları tamamen hatalıdır. | Konuşma esnasında vurgu ve tonlamaları kısmen hatalıdır. | Konuşma esnasında vurgu ve tonlamaları iyidir. |
| | Jest ve mimikler | Konuşurken konuşmasını destekleyecek jest ve mimiklere yer vermemiştir. | Konuşurken konuşmasını destekleyecek jest ve mimiklere kısmen yer vermiştir. | Konuşurken konuşmasını destekleyecek jest ve mimiklere yer vermiştir. |
| | Konuşma Hızı | Konuşmanın doğal akışını bozacak ölçüde yavaş konuşmaktadır. | Konuşmanın doğal akışını kısmen bozacak ölçüde yavaş konuşmaktadır. | Konuşma hızı normaldir. |

Additional 2

Yabancı Dil Olarak Türkçe Öğrenenler için A2 Düzeyi Karşılıklı Konuşma Becerisi Dereceli Puanlama Anahtarı

| | Maddeler | 1 | 2 | 3 |
|------------------------------------|-------------------------------|--|--|---|
| Konuşmanın İçeriği | Başlangıç | Konuşmaya uygun ifadelerle başlamamıştır. | Konuşmaya kısmen uygun ifadelerle başlamıştır. | Konuşmaya uygun ifadelerle başlamıştır. |
| | Ana Düşünce | Konuşmasında ana düşünceye yer vermemiştir. | Konuşmasında ana düşünceye kısmen yer vermiştir. | Konuşmasında ana düşünceye yer vermiştir. |
| | Bitiş | Konuşmayı sonlandıracak ifadeleri kullanmamıştır. | Konuşmayı sonlandıracak ifadeleri kısmen kullanmıştır. | Konuşmayı sonlandıracak ifadeleri kullanmıştır. |
| | Konuşma Esnası | Konuşma esnasında belirgin problemler yaşamıştır ve konuşmanın karşılıklı akışı bozulmuştur. | Konuşma esnasında daha az problem yaşamıştır ve konuşmanın karşılıklı akışı kısmen etkilenmiştir. | Konuşma esnasında problem yaşamamıştır ve konuşmanın karşılıklı akışı devam etmiştir. |
| Söz Varlığı | Kelime Hazinesi Durumu | Kelime hazinesi yetersizdir. | Kelime hazinesi sınırlıdır. | Kelime hazinesi yeterlidir. |
| Konuşmanın Şekilsel Boyutu | Dilbilgisel Hatalar | Konuşurken hatalı dilbilgisel yapılar kullanmıştır (Yaptığı konuşmasının anlaşılmasını etkilemektedir.) | Konuşurken hatalı dilbilgisel yapılar kullanması konuşmasının anlaşılmasını kısmen etkilemektedir. (Konuşurken dil bilgisel hataları vardır ancak ne söylemek istediği anlaşılmaktadır.) | Konuşurken hatalı dilbilgisel yapılar kullanımı yok denecek kadar azdır. (Konuşurken anlaşılmasını etkileyecek dil bilgisel hatalar yapmamıştır.) |
| | Akıcılık | Konuşması akıcı değildir. Çok fazla gereksiz duraklamalar ve kelime seçimlerinde tereddütler yaşamıştır. | Konuşması kısmen akıcıdır. Kısmen daha az duraklama ve tereddüt yaşamıştır. | Konuşması akıcıdır. Konuşurken gereksiz duraklamalar yapmamıştır. |
| | Telaffuz | Telaffuz hataları çok fazladır. Telaffuzu anlaşılır değildir. | Telaffuz hataları göreceli daha azdır. Telaffuzu kısmen anlaşılırdır. | Telaffuz hataları yok denecek kadar azdır. Telaffuzu anlaşılır ölçüdedir. |
| | Vurgu ve Tonlama | Konuşma esnasında vurgu ve tonlamaları tamamen hatalıdır. | Konuşma esnasında vurgu ve tonlamaları kısmen hatalıdır. | Konuşma esnasında vurgu ve tonlamaları iyidir. |
| | Jest ve mimikler | Konuşurken konuşmasını destekleyecek jest ve mimiklere yer vermemiştir. | Konuşurken konuşmasını destekleyecek jest ve mimiklere kısmen yer vermiştir. | Konuşurken konuşmasını destekleyecek jest ve mimiklere yer vermiştir. |
| | Konuşma Hızı | Konuşmanın doğal akışını bozacak ölçüde yavaş konuşmaktadır. | Konuşmanın doğal akışını kısmen bozacak ölçüde yavaş konuşmaktadır. | Konuşma hızı normaldir. |
| Konuşmayı devam ettirebilme | İletişimi sürdürebilme | Sorulan soruyu anlamamıştır ve karşılıklı konuşmak için yeterli alıcı dil becerilerine sahip değildir. | Sorulan soruyu anlamıştır ve karşılıklı konuşmayı sürdürebilmek için kısmen yeterli bir konuşma performansı göstermiştir. | Sorulan soruyu anlamıştır ve karşılıklı konuşmak için yeterli performansı göstermiştir. |