Assessment and peer assessment of foreign speaking production competence of the 1ˢᵗ year pre – service teachers

D. Stelmakh

Kyiv National Linguistic University
Corresonding author. E-mail: distelmah@gmail.com, https://orcid.org/0000-0001-5537-6740

Abstract. The article focuses on the assessment and peer – assessment of the foreign language speaking production skills of the 1ˢᵗ year pre – service teachers. The analysis of the existing evaluative tasks for speaking assessment has been done and best tasks for 2 types of monologues – a monologue – description and a monologue – narration have been collected. Successful assessment requires definite criteria to be chosen. In this paper criteria have been established and explained and 2 rating scales for a monologue – description and a monologue – narration have been developed.

Keywords: assessment, peer – assessment, foreign speaking production, pre – service teachers.

Introduction. Foreign language speaking assessment has always been in the center of the scholars' attention. Assessing foreign language speaking production competence is an important part of assessing speaking. On the one hand a wide range of skills required for different types of speaking production makes the process of assessing foreign language monologues rather difficult. On the other hand, according to the demand of modern approaches (reflexive, professional) assessment skills are essential for the professional competence of the pre – service teachers and therefore have to be developed.

Literature review. Foreign language speaking assessment has become the purpose of different researchers all over the world. Various aspects of foreign language speaking assessment have been discussed. Some of the researches focused on the process of evaluating speaking in general (Brown & Abeywickrama, Fulcher, Kitao Knight, Luoma, Thornbury, Ukrainska etc.). Some tried to find out how distinct the levels of second language proficiency are (Iwashita, et al). A lot of scholars focused on descriptors and rating scales (Fairbairn & Dunlea, Galaczi, et al, Kvasova, Sandrock et al). Research of peer assessment and self – assessment of speaking skills has also aroused the interest (Babaii et al, Cheng & Warren, Joo, Musfirah).

The purpose of this article is to choose the most optimal assessment tasks, to establish number of criteria and develop rating scales for 2 types of the monologues (a monologue – description and a monologue – narration (including a monologue – giving information) with the aim of assessment and peer – assessment of productive speaking skills of the first – year pre – service teachers.

Materials and methods. To achieve the purpose of the research number of general scientific methods such as analysis, synthesis, generalization, comparison have been used.

Results and their discussion. Assessment speaking skills of pre –service teachers is an essential part of the educational process in Ukrainian universities. For historical reasons foreign language interaction and production skills on the territory of Ukraine have been evaluated separately for a long time and this tradition still exists. According to the Educational Curriculum pre service teachers must develop their foreign language speaking production skills in different types of monologues, hence system of evaluation requires different tasks for different monologues. What is more, as Luoma (2004) with reference to Bygate states that if someone is good at one type, that does not mean he can be good

at the others (p.32). According to the communicative purpose Ukrainian scholars differentiate three main types of the monologues: a monologue – description, a monologue narration, which includes a monologue – giving information and argumentative or persuasive monologues (Ustimenko, 2013). In our research we focus on a monologue – description and a monologue – narration (including a monologue – giving information) as argumentative or persuasive monologues are more complicated and mostly used in assessment practice during the senior courses. In addition, they can be rather difficult for peer assessment. To solve the 1ˢᵗ task of this research it is necessary to define what speaking tasks are and to study the range of the tasks offered by different scientists for measuring foreign language speaking skills. Luoma (2004), modifying the definition of Bachman and Palmer explains speaking tasks as 'activities that involve speakers in using language for the purpose of achieving a particular goal or objective in a particular speaking situation' (p.31). As our aim is to choose the best ones for particular types of monologues it is necessary to study those which are suitable for narration, description or both. Firstly, let's study descriptive tasks. Description is frequently used in different exams and one of the typical tasks is *description of any familiar object or a person of a test –taker (a flat, a room, a car, a university, a friend, a relative etc.).* The advantages of such tasks, in our opinion, are a real life situation (students produce monologues basing on real objects or people from their life) and a wide range of topics which could be included. Among the other pluses are possibilities for direct and distant assessment (Luoma, 2004). Description often bases on some visuals as visual stimuli are an economical and effective way of introducing a situation or a topic of conversation without providing the candidate with a lot of vocabulary (Underhill, 1987) The next task which can be suitable for descriptive monologues is *describing a picture/ a photo.* This task – type gives the opportunity to the assessor check how detailed the speaker is though on the other hand it can be less communicative if the raters have already seen the picture described. What is more, it is necessary to remember that pictures or photos used for the assessment of the monologues must meet some requirements. Ukrainska (2009), particularly, outlines the following demands for the images: rather big size of pictures or photos, not more than 6 pictures or photos for 1 task, vivid and high quality images (black or white), informative pictures which can provide necessary length of oral production, topics or situations of pictures or

photos must match the Curriculum, they also must depict situations which are easy to recognize and be interesting to the test – takers (p.45). *Comparing pictures or photos* is another variation of the previous task which gives the opportunity for the assessor to check not only descriptive skills of the examinee but also his skills of comparison and contrast.

The second group of tasks include those which are suitable for narrative monologues. *Telling a story from your life* is one of such tasks. This task satisfies the communicative function of narration but Luoma (2004) says it can be "difficult to replicate in a test situation", because speaker doesn't always want to share some personal details. (p. 144). Such tasks can be used for peer assessment as it is much easier to share details with a friend than with a teacher. *Telling a picture – based story* is another task which is suitable for the narration. The advantage of it is that the test – taker does not have to tell any personal details and recollect some stories from his experience, the context is provided by the pictures (usually series of images is used for the narration). *Explaining the graph or chart* is a good task for a monologue – giving information. Usually graphs represent some information which is necessary to explain, so it is a good stimulus for speaking, but it is necessary to take into account the complexity of them and remember that graphs mustn't be too difficult for the examinees and meet the requirements for the images mentioned above. *A story paraphrasing/a story retelling* are also common tasks for checking students' speaking but they are mostly used during formative assessment as such tasks are considered by some scientists as reproductive ones which are not suitable for the communicative approach (Ukrainska, 2009).

The third group of tasks for checking foreign speaking production comprises those which are appropriate for both narrative and descriptive monologues. Simulated tasks are widely used for different types of speaking production. For instance, role- play. *Role – play* is an assessment task in which students play certain role basing on a typical or imaginary situation (Ockey & Li, 2015) There are a lot of benefits of such tasks among which are the following: it helps to develop students' creativity; it gives the opportunity to the test taker to elicit discourse that otherwise can be difficult to elicit; role – play lets students approach real – world situations (Sulispriyanto & Setyawan. 2015). *A speech basing on the situation* is a variation of a role – play, but the speaker usually speaks from himself, though sometimes can receive a role (Ukrainska, 2009) *An oral presentation or a report* is a good way to check both descriptive and narrative monologues. It is especially valid in an educational sphere (Thornbury, 2005). What is more, presentations can be uploaded to sites like You tube where teachers and groupmates can watch and comment them (Yimenes & Carrasco, 2013). One of the disadvantages of such task is that it definitely requires some preparation so speech spontaneity can't be checked by it. On the other hand, if a test – taker is asked questions after the presentation it increases the reliability of this task.

The second aim of this research was to establish the criteria for assessment of the narrative and descriptive monologues. There is a wide range of criteria offered for speaking assessment in different papers. Among them we can meet fluency, intonation, rhythm, pronunciation, vocabulary, grammar, coherence, discourse management, style, speech tempo, number of simple and complex sentences, duration of speaking etc. It is evident that using all of them can make oral production assessment practice rather sophisticated and time – consuming and therefore we have either to choose the most appropriate ones or to combine some of them. While deciding on the best criteria it is recommended not to choose more than 4 -5 features of the monologues (Ukrainska, 2009).

In Ukrainian universities speaking skills have been long evaluated from the perspective of the language accuracy, though nowadays the most important in speaking interaction and speaking production is a communicative purpose. There are a lot of errors which do not prevent the speakers from mutual understanding (Molokovych, 2001). When test – takers deal with descriptive or narrative monologues they aim at informing, describing or telling something to the listener, expressing or explaining his/her opinion. If the speaker managed to do that, the communication is successful. As a result, the first qualitative criterion we single out is the *realization of the communicative purpose.*

Despite the fact that *language accuracy* is considered to be a tool for the communication it influences the last one greatly. If the speakers' grammar or vocabulary skills are poor or limited it is quite difficult to reach the aim of communication. What is more a long – turn oral production requires a wide range of grammar or vocabulary to keep the listener involved. Language accuracy in our research also includes appropriate use of phonology, as clear pronunciation as well as intonation firstly allow to understand the information and, secondly, are especially important for future teachers of English.

The next parameter which can help measure the oral production is *content and coherence. Content* implies appropriateness of the monologue to the situation and to the discourse type and reflects the ability of the speaker to develop the topic. *Coherence* shows how logically and clearly the oral discourse is organized (Fairbairn & Dunlea, 2017).

Fluency has been the criterion for speaking assessment for a long time (Fulcher, 2003). De Jong (2018) in his study of fluency gives a definition of it from the glossary: fluency is "the flow in spoken or written language as perceived by the listener or reader" (p. 238). Fluency usually deals with speech rate, number of filled and silent pauses, number of repetitions and restarts etc. (De Jong, 2018) When judging students' fluency assessors usually evaluate the ability to speak naturally without worrying too much about making mistakes (Ulker, 2017).

Creating rating scales for narrative and descriptive monologues for pre – service teachers is the next step in assessment procedure after defining the criteria. There are 2 types of rating scales holistic and analytic: holistic base on giving a grade on the overall impression and analytic scoring give a separate mark for a separate aspect (Thornbury, 2005). While using holistic ones it is necessary to single out the levels of the students' oral performance and evaluate according to them. Analytic scales demand the description of the level basing on each criterion. As our scales are both teacher and student oriented it is more reasonable to use the analytic ones, because sometimes students hesitate which level to choose and, what is more, getting acquainted with the criteria is a professional necessity for the pre – service teachers.

A monologue – description assessment scale

Realization of the communicative purpose	Language range and accuracy	Coherence	Fluency
Level A (100 – 90 points)			
Communicative purpose is fully realized. A speaker gives full and detailed description naming main and additional characteristics so the listeners can easily visualize the described object	Language range and accuracy is excellent. A speaker uses a wide range of vocabulary and grammar. A lot of descriptive adjectives, idiomatic phrases, various grammatical structures of comparison and contrast, fully appropriate examples and justifications are used. Figurative language can be used. Speaker's sounds are distinct and intonation patterns are diverse. Mistakes hardly ever happen.	The structure of the description is logical. A speaker gives main and additional characteristics of the object one after another supporting each of them with examples or justifications and often comparing or contrasting the object with the opposite one. Fully appropriate cohesive devices are used.	The description is fluent. There can happen 1 - 2 hesitations, pauses or self-corrections, which do not last long.
Level B (89 – 83 points)			
Communicative purpose is mostly realized. A speaker gives detailed description of the object without naming some of the additional characteristics which do not prevent the listeners from visualizing the object.	Language range and accuracy is very good. A speaker uses a wide range of most frequent vocabulary and grammar. Descriptive adjectives, idiomatic phrases, grammatical structures of comparison and contrast, good examples and justifications are often used. Speaker's sounds are distinct and intonation patterns are rather diverse. Mistakes sometimes happen but they do not prevent understanding.	The structure of the description is logical. A speaker gives most important characteristics of the object one after another supporting each of them with examples or justifications sometimes comparing it with the opposite one. Appropriate cohesive devices are used.	The description is rather fluent. Some hesitations, pauses or self-corrections happen, but they do not last long
Level C (82 – 75 points)			
Communicative purpose is realized. A speaker gives description of the object giving basic characteristics of it so listeners can visualize it in general.	Language range and accuracy is good. A speaker uses enough vocabulary and grammar to describe the object. Descriptive adjectives, grammatical structures of comparison and contrast, examples and justifications are sometimes used. Pronunciation and intonation are sometimes incorrect. Grammar mistakes often happen.	The structure of the description is mostly logical. A speaker gives basic characteristics of the object supporting it with examples or justifications but a speaker uses only frequently occurring cohesive devices to describe the object as a list of basic characteristics	The description is fluent when simple sentences are used. Complex sentences cause a lot of hesitations, pauses or self – corrections.
Level D (74 – 68 points)			
Communicative purpose is partly realized. A speaker gives only several main characteristics of the object, so it is hard to visualize it.	Language range and accuracy are poor. A speaker uses simple vocabulary and grammar structures for description. Examples are rarely used. Speaker's pronunciation and intonation are often incorrect and causes some difficulties for the listener. Only simple sentences are grammatically correct and some grammar mistakes can lead to misunderstanding.	The structure of the description is not always logical. A speaker gives one characteristic but do not develop it and goes to the other. Then may return to the previous one. A set of cohesive devices is limited.	The description is slow with a lot of hesitations, pauses and self – corrections.
Level E (67 – 60 points)			
Communicative purpose is hardly realized. A speaker gives a few characteristics of the object. It is impossible to visualize it.	Language range and accuracy are very poor. A speaker uses very simple memorized vocabulary and grammar structures. Examples are hardly ever used. Speaker's pronunciation and intonation are often incorrect and cause a lot of difficulties for the listener. Grammar mistakes happen in almost all the sentences and often lead to misunderstanding.	The structure of the description is not logical. A speaker gives a few characteristics in disorder, do not develop them, often repeats some ideas. A set of cohesive devices is very limited.	The description is very slow with a lot of hesitations, very long pauses and repetitions. Self – corrections rarely happen, cause a speaker does not notice mistakes.
Level FX (59 – 35 points)			
Communicative purpose is not realized. A speaker gives very few unclear characteristics of the object. It is impossible to visualize it.	Language range and accuracy are so poor that it is impossible to evaluate the description. A speaker uses very few memorized words and grammar structures. Speaker's pronunciation and intonation are impossible to understand.	The structure of the description is not logical. A speaker gives very few characteristics in disorder. Phrases are mostly isolated. A set of cohesive devices is very limited	The description is so slow that pauses last longer than phrases or words.

According to the Curriculum at the end of the 1st course pre – service teachers' oral productive skills must satisfy level B2 of Common European Framework of References (CEFR) (Curriculum, 2001). Therefore, B2 will be the upper level of our scales. As Ukrainian students in the university study according to credit – modular system and are assessed within levels from A to F our scales will include 6 levels totally. It is also necessary to admit that all the criteria in the scales will be equally important as we believe that each of them reflects one side of the oral production competence. Before representing the scale for a monologue – description, let us analyze the micro skills of it which, according to Molokovych (2001), include the following skills: to develop and justify own thoughts sustaining logical order in a monologue – description; to speak coherently on the basis of several sources of information with the support of a given example or without it; to frame out own speaking relatively correct and according to the principle of a communicative adequacy with the means of the studied language (p. 46 – 47).

The second scale will represent levels and criteria appropriate for the assessment of a monologue – narration, the micro skills of which include the following ones: to tell about something (about yourself, your friend, your city\village etc.) according to the communicative situation, expressing own thoughts and own estimation; to develop and justify own thoughts in a chronological order; to finish speaking according to the content and structure (Molokovych, 2001, p.46).

A monologue – narration scale

Realization of the communicative purpose	Language range and accuracy	Coherence	Fluency
Level A (100 – 90 points)			
Communicative purpose is fully realized. A speaker fully and completely tells\transmits a story\information unfolding main and additional facts, characters and events so the listeners can easily retell it in details.	Language range and accuracy is excellent. A speaker uses a wide range of vocabulary and grammar. Idiomatic and figurative vocabulary can be used. A variety of tenses in active and passive voice, different clauses, adverbs of time, manner and place and descriptive adjectives are used. Speaker's sounds are distinct and intonation patterns are diverse. Mistakes hardly ever happen.	The structure of the narration is logical. A speaker tells\transmits a story\information in a chronological order, sequence of events is absolutely clear, spatial and temporal orientations, cause and effect relations are present. A story\information includes an introduction, main part and conclusion. Fully appropriate cohesive devices are used.	The narration is fluent. There can happen 1 - 2 hesitations, pauses or self-corrections, which do not last long.
Level B (89 – 83 points)			
Communicative purpose is mostly realized. A speaker fully and completely tells\transmits a story\information unfolding main and additional facts, characters and events without some unimportant details so the listener can easily retell it.	Language range and accuracy is very good. A speaker uses a wide range of most frequent vocabulary and grammar. A variety of tenses in active and passive voice, different clauses, adverbs of time, manner and place and descriptive adjectives are often used. Speaker's sounds are distinct and intonation patterns are rather diverse. Mistakes sometimes happen but they do not prevent understanding.	The structure of the narration is logical. A speaker tells\transmits a story\information in a chronological order, sequence of events is clear, spatial and temporal orientations, cause and effect relations are present. A story \information includes an introduction, main part and conclusion. Appropriate cohesive devices are used.	The narration is rather fluent. Some hesitations, pauses or self-corrections happen, but they do not last long
Level C (82 – 75 points)			
Communicative purpose is realized. A speaker tells\transmits a story\information unfolding all main facts characters and events without naming additional ones so the listener can retell it in general.	Language range and accuracy is good. A speaker uses enough vocabulary and grammar to unfold the main line of a story\information. Basic tenses mostly in active voice, rarely clauses, adverbs of time, manner and place and descriptive adjectives are sometimes used. Pronunciation and intonation are sometimes incorrect. Grammar mistakes often happen in complex sentences.	The structure of the narration is mostly logical. A speaker tells\transmits a story\information in a chronological order though sometimes the sequence of events requires explication. Spatial and temporal orientations, cause and effect relations are present but not enough to understand the story \information in details. There is a short introduction and a short conclusion in a story \information, a main part is not detailed. A speaker uses only frequently occurring cohesive devices to tell\transmit a main line of a story\information	The narration is fluent when simple sentences are used. Complex sentences cause a lot of hesitations, pauses or self – corrections.
Level D (74 – 68 points)			
Communicative purpose is partly realized. A speaker tells\transmits a story\information unfolding only several main facts, characters and events and it can be difficult for the listener to retell some parts of it.	Language range and accuracy are poor. A speaker uses simple vocabulary and grammar structures to tell\transmit a story\information. Basic tenses in active voice are used. A narration includes limited number of adverbs and adjectives Speaker's pronunciation and intonation are often incorrect and causes some difficulties for the listener. Grammar mistakes often happen in all types of sentences and can lead to misunderstanding.	The structure of the narration is not always logical. A speaker tells\transmits a story\information breaking the chronological order so it's difficult to understand the sequence of events. Spatial and temporal orientations, cause and effect relations are often omitted. The speech does not include distinct parts. A set of cohesive devices is limited.	The narration is slow with a lot of hesitations, pauses and self – corrections.
Level E (67 – 60 points)			
Communicative purpose is hardly realized. A speaker tells\transmits a story\information unfolding a few facts, characters and events so it is difficult for the listener to understand and retell it.	Language range and accuracy are very poor. A speaker uses very simple memorized vocabulary and grammar structures. One or 2 grammar tenses in active voice are used. A narration includes several adverbs and adjectives and a speaker often repeats them. Speaker's pronunciation and intonation are often incorrect and cause a lot of difficulties for the listener. Grammar mistakes happen in almost all the sentences and often lead to misunderstanding.	The structure of the narration is not logical. A speaker tells\transmits a story\information without any chronological order so it's very difficult to understand the main idea. There are a few spatial and temporal orientations, cause and effect relations. The narration does not include distinct parts. A set of cohesive devices is very limited.	The narration is very slow with a lot of hesitations, very long pauses and repetitions. Self – corrections rarely happen, because a speaker does not notice mistakes.
Level FX (59 – 35 points)			
Communicative purpose is not realized. A speaker tells\transmits a story\information unfolding very few facts, characters and events so it is impossible to understand and retell it.	Language range and accuracy are so poor that it is impossible to evaluate the narration. A speaker uses very few memorized words and grammar structures. Speaker's pronunciation and intonation are impossible to understand.	The structure of the narration is not logical. A speaker tells\transmits a story\information without any chronological order so it's impossible to understand the main idea. There are no spatial and temporal orientations and cause and effect relations. The narration does not include distinct parts and phrases are mostly isolated. A set of cohesive devices is very limited	The narration is so slow that pauses last longer than phrases or words.

Conclusions. In this article an attempt to improve the process of assessment and peer – assessment of the foreign language speaking production skills of pre-service teachers was made. Particularly, tasks for evaluation of a monologue – description and a monologue narration have been

collected, criteria have been established and analytical rating scales have been created. The further research will base on the development of a general rating scale for the assessment and peer – assessment of foreign language production skills of pre – service teachers.

REFERENCES

1. De Jong, N. H. (2018) Fluency in Second Language Testing:Insights From Different Disciplines. *Language Assessment Quarterly*, 15:3, 237-254.
2. Fairbairn, J., Dunlea, J. (2017). *Technical report. Speaking and writing scales revision.* British council.
3. Fulcher, G. (2003). *Testing second language speaking*. Harlow: Longman/Pearson Education.
4. Luoma, S. (2004). *Assessing speaking.* Cambridge: Cambridge University Press.
5. Molokovych, O. O. (2001). *Testing Speaking of the Second Year Students Studying English as a Second Foreign Language.* Kyiv national linguistic university. Kyiv.
6. Nikolayeva, S.Yu., Solovei, M. I., Golovach, Yu. V., Inozemtseva, I. O., Kamynin I. M., Kolominova, O. O. …Tyaglovska V. M. (2001).Curriculum for English language development in universities and institutes. Kyiv.
7. Ockey, G. J., Li, Zhi. (2015). New and not so new methods of assessing oral communication. *Language value,* 7, 1 – 21.
8. OECD (2021), *PISA 2025 Foreign Language Assessment Framework*, PISA, OECD Publishing. Paris.
9. Sulispriyanto, E., Setyawan, FX. O. (2015) *Assessing speaking. A brief summary.* Universitas Slamet Riyadi.
10. Thornbury, S. (2005). How to teach speaking. Pearson – Longman.
11. Ukrainska, O. O. (2009) *Methodology of testing competence in speaking French of pre – service teachers.* Kyiv national linguistic university. Kyiv.
12. Ulker, V. (2017). The design and use of speaking assessment rubrics. *Journal of education and practice,* 8, 135 – 141.
13. Ustymenko, O. M. (2013). Formation of foreign language monologue competence. In S. U. Nikolayeva (Ed.), *Methodology of teaching foreign languages and cultures: theory and practice* (pp.340 – 368). Kyiv: Lenvit.
14. Yimenes, B. A., Carrasco, Y. N. (2013). Alternative Assessment: A Feasible Way to Assess Oral Production Through ICTs. Valdiviaa.