sciendo International Conference KNOWLEDGE-BASED ORGANIZATION
Vol. XXVIII No 2 2022

STANAG 6001 RUSSIAN LANGUAGE EXAM: HOLISTIC APPROACH TO ASSESSING SPEAKING

Diana BORIMECHKOVA

"G.S. Rakovski" National Defence College, Sofia, Bulgaria
d.borimechkova@rndc.bg

Abstract: *This paper analyzes the STANAG 6001 Russian language exam focusing on the specifics of Speaking Proficiency evaluation. Despite the fact that the testing procedure is based on the Assessment Criteria set by the STANAG 6001, assessing speaking skills through a guided interview or a sequence of techniques seems to be one of the most challenging and hard tasks for testers. Advantages of holistic approach of assessing speaking skills are discussed and justified.*

Keywords: STANAG 6001 testing, Russian language, speaking assessment, holistic approach

1. Introduction

It is usually assumed that the Standardization Agreement STANAG 6001 refers only to the evaluation of English language proficiency of the military personnel of NATO member countries. However, contemporary geopolitical situation requires that the military and civilian personnel more than ever keep high level of foreign language skills and cultural awareness in order to face the new challenges. That is why the proficiency in second or even third foreign language is a great advantage and is ever valued for everyone who takes part in NATO exercises, peacekeeping missions, international trainings or work in multinational staff. STANAG 6001 tests are the official exams used in NATO for employment and deployment decisions, promotions, etc. The primary purpose of all STANAG 6001 language proficiency tests is evaluation of the candidate's ability to communicate in a foreign language for the purpose of interoperability within NATO and in order to perform specific professional tasks.

Given the increased interest in Russian language learning and testing, Rakovski National Defence College (RNDC) in Sofia is the only Bulgarian institution administrating Russian language tests according to NATO Standardization Agreement STANAG 6001. Russian STANAG 6001 is a ***multilevel test*** covering levels 1, 2 and 3 aimed to assess the proficiency of candidates in four language skills – listening, reading, writing and speaking.

In comparison with the relatively easy score-based evaluation of multiple-choice reading and listening tests, assessing productive skills – writing and speaking in particular – seems to be one of the most challenging and hard-to-achieve-objectivity tasks because of the numerous factors affecting the test structure, assessment procedure, perception and possible subjectivity of the assessors, timing factors to name a few.

DOI: 10.2478/kbo-2022-0059
© 2015. This work is licensed under the Creative Commons Attribution-Non Commercial-No Derivatives 3.0 License.

Russian language speaking test is based on a guided interview comprising different speaking tasks (narration, description, giving instructions, hypothesizing, etc.) during which the testers should take into account the wide range of competences that candidates strive to demonstrate and prove. The focus of the research is the Speaking component of STANAG 6001 Russian Language Exam and a special attention to the speaking assessment. Thus, the goal of this paper is to underline the advantages of holistic approach in comparison with the drawbacks of formal analytic approach to assessing speaking performance of the candidates.

2. General characteristics of the STANAG 6001 test

Standardization Agreement "Language Proficiency Levels" (STANAG 6001) Edition 5 of December 11, 2014 is intended for NATO military and civilian personnel who are non-native speakers with the aim "to be used as the common standard for language curriculum and test development, for recording and reporting Standardised Language Profiles (SLPs)" [1]. STANAG 6001, like the Common European Framework of Reference for Languages (CEFR), can be defined as a standard for describing language proficiency and it is not based on any curriculum, language programme or language course book [2]. Testing process within the STANAG 6001 is carried out in accordance with the rules and unique assessment criteria established by the Bureau for International Language Coordination (BILC), which is the NATO's consultative and advisory body for language training and testing. This study was provoked by the process of equating the Bulgarian National Test Specifications of English with the Russian specifications and making them correspond to the BILC documents [3]. The most important characteristics of STANAG 6001 testing are summarized in the table below.

Table 1 General Characteristics of STANAG 6001 Tests

STANAG 6001 tests are:	
• **PROFICIENCY**	*Assess an individual's ability to consistently complete the real-world communication tasks for everyday communicative needs and in a military-related context.*
• **FOUR-SKILLS**	Reading, Listening, Writing and Speaking *Each skill is assessed independently of the other skills.*
• **MULTILEVEL**	for Levels 1, 2 and 3 *Each higher level includes all the language abilities of the lower level.* NB: Apart from the multilevel testing since 2017, Bulgaria has been implemented a STANAG 6001 English Language Test for Level 4.
• **CRITERION-REFERENCED**	Scales and scores reflect language competence defined in the simplified level description of the Standardized Language Profiles (SLPs) [3] as follows: Level 0 – No proficiency Level 1 - Survival Level 2 - Functional Level 3 – Professional

The following two interrelated distinctive features distinguish STANAG 6001 from other language test formats:

- *Lack of the component 'Language in use'*

Direct testing of grammar, vocabulary or language in use is not included in STANAG 6001 text format, because the linguistic component is implicitly included and assessed in tests of reading, listening, writing and speaking.

- *Military flavour*

It is important to emphasize the pragmatic orientation of STANAG 6001 exam format with its main purpose to evaluate the degree of the candidate's ability to perform real-world tasks in specified situations. Characteristics inherent to language for special purposes proficiency – military-oriented topics and tasks – are recommended by BILC and included into detailed descriptors of STANAG 6001 Test Specifications of all languages [4], [5].

3. Russian STANAG 6001 test

Russian language exam is intended for Bulgarian Armed Forces military and civilian personnel who are non-native speakers of Russian. The test gives assessment of the extent to which a candidate is able to "communicate effectively in everyday situations and professional environment in order to participate in multinational military exercises, operations, and missions". Target language use situation is broadly defined as "Russian for everyday communicative needs and in a military-related context" [4]. Since 2020, Russian Language STANAG 6001 test has been updated and the 2nd revised edition of Russian Language Test Specifications has come into force in order to implement multilevel test, to standardize the types of tasks for all skills and to meet BILC recommendations on introducing assessment criteria and 'plus' levels in defining the Standardized Language Profile (SLPs) of the candidates. Test Specifications and the test materials have been designed in full accordance with the theoretical and methodological language testing principles and descriptors of the language proficiency levels of STANAG 6001, Edition 5. New items and test variants are constantly being developed by the Exam Board comprising two experts of the Language Training Department at RNDC. It is worth mentioning that designing materials and defining the procedures for proficiency test which assesses the four language skills requires a lot of expertise and rigorous work. The Russian language testing is conducted at RNDC in two exam sessions per year. During the period 2017-2022 up to 30-40 candidates have been tested per year in accordance with STANAG 6001.

4. Testing Speaking according to STANAG 6001: the Russian case

According to the generally accepted opinion of testers, test-takers and stakeholders, speaking requires communication in real time and may be one of the most challenging to master, yet it is most valuable of the four skills. The candidate is expected to demonstrate effective communicative use of the target language through various discourse patterns and strategies, appropriate grammatical structures and vocabulary, register and style relevant to the context [3].

STANAG 6001 speaking test consists of a variety of language tasks on a range of topics. According to the STANAG 6001 descriptors on Speaking, *Topics* covered for Levels 1 and 2 can be drawn from the personal and public domains, including daily routines, giving instructions, descriptions, free time activities, etc., and for Level 3 professional and military issues should be included, i.e. international relations, defence and security issues, social life, science and technologies, ecology, etc. The length of the Russian speaking test is up to 30 min depending on the level of the candidate [4]. The candidate and a board of two qualified interviewers are involved.

The stages of the Russian speaking test are as follows:

- Warm-up and Introduction: self-presentation of the candidate;
- Series of increasingly complex questions and topics to establish the level;
- Questions and Answers: general conversation to elicit candidate's opinion;
- Discussion and Conversation on abstract topics, including military ones;
- Wind-down and feedback.

The procedure consists of a mix of techniques: guided interview; role-play, and discussion/conversation. The interview is a direct face-to-face interaction between the interviewer/s and the test-taker. It is considered the most authentic assessment tool because though it follows a pre-determined structure and protocol and reflects a real-life conversation with a certain degree of flexibility and 'freedom'. Using a sequence of speaking test techniques allows the interviewer to get the best of the candidate in the shortest possible time. Summarizing the above, the Russian speaking test is:

- multi-level: it measures proficiency from Level 1 to Level 3 in accordance with 7 descriptors for levels 0, 0+, 1, 1+, 2, 2+ and 3;
- adaptive: tasks vary depending on the candidate's language proficiency;
- interactive: involves actively both parts – the interviewer/s and the test-taker;
- based on a guided interview and a sequence of techniques during which the exam board should take into account the wide range of competences;
- based on rating scales.

5. Assessment Criteria used for Speaking STANAG 6001

Speaking is the most important marker of foreign language proficiency that differs from all other skills by involving a great number of extralinguistic factors concerning candidates' performance or test reliability. According to both English and Russian STANAG 6001 Test Specifications, the following criteria should be taken into account in the speaking assessment: communicative skills, control of vocabulary and grammatical structures, register, style, pronunciation, and fluency [4], [5].

Subjectivity in its most positive sense has always been and will remain an immanent feature of the speaking assessment. Intrinsically, speaking testing procedures and assessment criteria of STANAG 6001 are based on the model of communicative competences referring to the ability to use language to communicate successfully. According to the model proposed and further expanded by Canale and Swain in 1980s, communicative competence consists of four equivalent elements: linguistic, strategic, sociolinguistic, and discourse competence. The Common European Framework of Reference for Languages (CEFR) divides communicative language competences in linguistic, sociolinguistic and pragmatic. A number of contemporary studies continue to research the factors involved in the process of testing speaking from the point of view of interaction, cognitive strategies, process competence, sociocultural capacity, etc. The endeavor of assessing an integral range of competencies corresponds to the main objectives of STANAG testing of speaking.

STANAG 6001 Speaking Proficiency Assessment Criteria [6] is a reference grid which includes 4 columns with assessment criteria and 7 bands with short descriptors of levels 0 to 3, including plus levels (0+, 1+ and 2+). Descriptors indicated horizontally in the reference grid are "series of short description of different levels of language ability" or *rating scales* [7]. Developing a valid, reliable rating scale with consistency and continuity of descriptors is one of the main requirements for testing in general.

Questions arise when analyzing vertical concepts of the proposed by BILC reference grid represented by four criteria with brief comments: *Content* (What a candidate can talk about: topics, subject areas addressed,

settings), *Tasks* (What a candidate can do with the language tasks accomplished, attitudes expressed, tones conveyed), *Accuracy* (How well a person can use the language: acceptability, quality and correctness of message conveyed) and *Text Produced* (Length and organization of texts; discourse types) [6]. Obviously, these criteria cover a very wide range of competencies and the recommendations of BILC do not offer a specific method of weighting different criteria. Are these criteria enough to evaluate the ability of the candidate to perform specific professional tasks in a military context? How can examiners actually use this scale? What is the real practice?

6. Holistic and Analytic Approaches to Speaking Assessment

There are two traditional evaluating methods or approaches: *holistic* and *analytic*. Holistic approach involves the assignment of a single score to a task on the basis of an overall impression of it. Analytic assessment requires assessing the candidates' proficiency based on multiple criteria awarding scores or marks for each one of them in order to establish an overall proficiency by totalling the criteria. It is generally assumed that analytic method boosts the reliability of the assessment [9]. Hence, it is more time-and-energy-consuming although it gives objective and abundant information on the language ability of a test taker [10]. Analytic assessment risks missing the whole when focusing on the detailed scores of the particular parts similarly to the way holistic assessment sometimes risks missing the details when focusing on the whole.

An analytic score-based assessment is usually applied to evaluate the oral performance in one-level testing formats, for example Cambridge English First Certificate (FCE) and Test of Russian as a Foreign Language (TORFL). However, we should not forget that the specifics of such tests are quite different from the STANAG 6001 format in their aims, tasks, scope, target groups, methodology and timing..

The Russian exam board assesses the performance of the candidate *as a whole* carefully considering the *rating scales* indicated in the level descriptors.

7. Advantages of the Holistic Assessment

The essential characteristics of STANAG 6001 testing suggest the use of a holistic approach to the speaking assessment which takes into account both linguistic and extralinguistic factors. By applying holistic approach, experienced interviewers are able to assess a number of additional categories that are problematic from the point of view of the quantitative analysis or measurement: communicative ability, topics addressed, tasks accomplished, fluency and naturalness of speech, self-confidence, use of cultural references and communication strategies to compensate for breakdowns in communication, etc.

The purpose of the assessment should insite the assessment approach. Proficiency tests for specific purposes "should be able to give a reliable indication of whether a candidate is proficient enough to carry out the tasks that will be required". Hence, holistic or *impressionistic* approach to the speaking performance is one of the possible and very often "preferable way to measure, in particular when it comes to the testing language for specific purposes" [7].

It is not possible to predict all the operations which will be performed in an interactive oral test, and that is why the perception of the interviewers and the impression made by the candidate play a very important role in this interview-based interactive testing procedure [10]. It is often said that the first impression is the best impression but – in our case – only if it made by a duly qualified interviewers who are at the same time assessors. The evaluation of the candidate is discussed and negotiated by both assessors who need to be well trained and aware of the level descriptions and assessment criteria, but also must have an idea of the job descriptions that require

foreign language proficiency.

As indicated in the main STANAG 6001 document, testing of language proficiency is an 'open' issue and all participant nations and NATO bodies are invited to "propose changes, at any time" [1]. That is why a further analysis of the assessment procedure, rating scales, and criteria is needed in order to better suit the NATO armies' current needs.

6. Conclusions

The examiners-assessors should take into account all the specifics and the purposes of the STANAG 6001 testing format and, based on pre-established criteria, should quickly take the decision on the level of competence of the candidates, i.e. determine the level of proficiency which includes not only language ability, but also strategic, sociolinguistic and pragmatic competencies and the ability to perform professional tasks in Russian. Therefore, given the nature of the multilevel, proficiency and professional-oriented exam, we believe that holistic assessment based on STANAG 6001 regulating documents, best practices, scientific research, clear criteria and the expert judgement of the examiners is the best way to assess the speaking skills.

References List

[1] NATO Standardization Office. Standardization Agreement "Language Proficiency Levels" (STANAG 6001) Edition 5. Available from: https://www.natobilc.org/

[2] Council of Europe. Common European Framework of Reference for Languages: Learning, teaching, assessment – Companion volume, Council of Europe. Strasbourg; 2020.

[3] BILC. Best Practices in STANAG 6001 testing, March 2019. Available from: https://www.natobilc.org.

[4] Specifications of the NATO STANAG 6001 Russian Language Proficiency Test for the Ministry of Defense of the Republic of Bulgaria in accordance with NATO STANAG 6001 ed. 5, 2nd revised edition. Available from: https//www.mod.bg

[5] Specifications of the NATO STANAG 6001 English Language Proficiency Test for the Ministry of Defense of the Republic of Bulgaria in accordance with NATO STANAG 6001 ed.4, 11th revised edition. Available from: https//www.mod.bg

[6] BILC. STANAG 6001 Speaking Proficiency Assessment Criteria. Available from: https://www.natobilc.org.

[7] Underhill N. Testing Spoken Language. A Handbook of Oral Testing Techiques. Cambridge University Press; 1987. pp. 88-100.

[8] Радкова А. Оценяване на писмената реч в чуждоезиковото обучение, Тих-Ивел, 2010. ISBN: 978-954-9310-21-4. pp. 141-160.

[9] Namaziandost E., Ahmadi Sh. The Assessment of Oral Profciency through Holistic and Analytic Techniques of Scoring: A Comparative Study. Applied Linguistics Research Journal, 3 (2): 2019. doi: 10.14744/alrj.2019.83792.

[10] Hughes A. Testing for Language Teachers 2nd edition. Cambridge University Press; 2003. pp. 102-105, 116-117.