ASSESSMENT LITERACY OF RUSSIAN AS A FOREIGN LANGUAGE (RFL) TEACHERS

Anastasia Drackert
Ruhr-University Bochum, Germany
anastasia.drackert@rub.de

Wolfgang Stadler
University Innsbruck, Austria
wolfgang.stadler@uibk.ac.at

КОМПЕТЕНЦИЯ ПРЕПОДАВАТЕЛЕЙ РКИ В ОБЛАСТИ ОЦЕНИВАНИЯ

Анастасия Ивановна Драккерт
Рурский Университет г. Бохума, Германия

Вольфганг Штадлер
Инсбрукский Университет, Австрия

ABSTRACT

The focus on competence-orientation in the Russian language classroom still confronts RFL teachers with new challenges not only when teaching but also in their assessment of students' performances. We report on the results of a needs analysis whose aim was to survey the assessment practices of in-service teachers of RFL (*N*=198) in Austria, Germany, Switzerland and South Tyrol (Italy) as well as their actual needs in terms of assessing student classroom performances in Russian. We conclude by presenting a list of recommendations for the training of pre-service and in-service RFL teachers.

АННОТАЦИЯ

Компетентностный подход в обучении иностранным языкам по-прежнему представляет особую сложность для преподавателей русского языка как иностранного (РКИ) не только при обучении, но и при оценке знаний учащихся. В статье представлены результаты исследования, цель которого заключалась в сборе данных о существующей практике, используемой преподавателями РКИ в Австрии, Германии, Швейцарии и в Южном Тироле (Италия) при оценивании навыков и умений владения русским языком у школьников, а также в определении их потребностей и проблем, возникающих при оценивании школьников. В заключении приводится список рекомендаций по подготовке и повышению квалификации преподавателей РКИ.

Keywords: Russian language classroom; assessment literacy; needs analysis; classroom-based assessment

Ключевые слова: Русский как иностранный в школе; компетенция оценивания; анализ потребностей; оценивание в классе

Introduction

An understanding of good assessment practices and the ability to apply this knowledge to assessing students, often referred to as 'assessment literacy', is fundamental to achieving and maintaining an overall good quality of teaching and learning. This recent acknowledgement of the central role that assessment plays in students' learning has led to a growing interest in assessment literacy in different educational contexts (Stiggins 1991). Since classroom teachers have been recognized as key agents in educational practice (Leung 2014; Bachman & Damböck 2017), it is only logical that assessment literacy has also become an important component of the professional competence of teachers in general and foreign language teachers in particular.

Language Assessment Literacy (LAL): definition and previous research

Language assessment literacy – understood as a specific type of assessment literacy – is defined as "the ability to design, develop and critically evaluate language tests and other assessment procedures, as well as the ability to monitor, evaluate, grade and score assessments on the basis of theoretical knowledge" (Vogt & Tsagari 2014: 377). In spite of two decades of research in LAL, only few studies have dealt with LAL of classroom teachers, some of them addressing their general LAL (Hasselgreen, Carlsen & Helness 2004; Vogt & Tsagari 2014), others focussing on language assessment literacy in a particular language skill, for example in writing (Crusan, Plakans & Gebril 2016). Among these studies, only one needs analysis has been conducted in Germany (Vogt & Tsagari 2014).

Vogt and Tsagari conducted a large-scale study in seven European countries (one of them being Germany) in order to explore perceived LAL levels of foreign language teachers and their training needs in language testing and assessment. The results of the German sample showed that the majority of the informants, namely teachers of English (EFL) in different educational institutions, had received no or little training in LAL. If training did take place, its focus was on testing micro-linguistic aspects and language skills. Innovative aspects of assessment, such as self-assessment or the use of the European Language Portfolio (ELP), were identified as important areas of professionalization. At the same time, EFL teachers expressed no need for training in the area of developing and marking classroom tests. Having compared the results from seven European countries, the researchers eventually concluded that professional training in assessment should be tailored to teachers' needs and should take into account their existing practices.

To find out the actual needs and existing assessment practices of RFL teachers – a group that has been neglected so far in LAL-research (Drackert & Stadler, 2017) – we carried out a survey among RFL teachers in four German-speaking countries.

RFL at secondary schools in German-speaking countries

In German-speaking countries, Russian is primarily taught as a second (former GDR), third or fourth foreign language (former West Germany and Austria), with the exception of Waldorf Schools, where Russian is frequently taught as the first foreign language. In many cases, there is only one RFL teacher employed at a school. Russian is considered a minor subject ('kleines Fach') despite the fact that 111,185 pupils were learning RFL in Germany in school year 2015/2016 and about 7,600 in Austria in school year 2014/15 (Source: *Tiroler Tageszeitung* 27.9.2017). Russian ranks fifth among foreign languages taught in secondary education (Mehlhorn 2016: 536).

For Russian there exist no comparative studies as there are for English or German (VERA[6] or DESI[7]), and there are only two standardized Russian tests (*TRKI, telc*) on the market, neither of which was developed for the secondary school learner population, but for adults. Furthermore, in comparison with English as a foreign language, there is hardly any literature including best-practice examples for assessing language skills in Russian. RFL teachers have only few opportunities to work together with colleagues or deal with external assessment, which has proved to be conducive not only to increasing the quality of tests but also to increasing the teachers' assessment literacy (Stadler & Kremmel 2018: 53). This deficit makes RFL teachers 'stew in their own juice', which has an influence not only on their assessment practices but also on their assessment literacy.

Study

Participants

A total of 198 RFL secondary school teachers took part in the survey: 109 from Germany, 65 from Austria, 17 from Switzerland und 7 from South Tyrol. Their teaching experience ranged from one to 30 years or more. Around 70% of the respondents are involved in conducting or assessing school-leaving exams in Russian. Most of them (*N*=141) teach Russian at grammar schools ('Gymnasien').

Data collection and analysis

A three-part online questionnaire developed by the authors was used for data collection. Part 1 included 12 questions on the teachers' professional background. A total of 13 questions in part 2 dealt with different aspects of assessment practices of RFL teachers with a focus on classroom assessment

---

6   VERA is short for *Vergleichsarbeiten*, which are written comparison tests that students take in the 3rd and 8th grades (VERA-3 and VERA-8) in Germany (cf. https://www.iqb.hu-berlin.de/vera).

7   DESI is short for *Deutsch-Englisch-Schülerleistungen-International (German-English-Classroom Performances-International)*. The study provides basic information on the performance levels of 9th grade students in English and their current mastery of the German language. It also endeavors to explain the differences in achievement levels and students' basic knowledge for educational and school policy intervention measures, the revision of curricula, textbooks and teaching materials as well as for teacher education (cf. https://www.kmk.org/themen/qualitaetssicherung-in-schulen/bildungsmonitoring/internationale-schulleistungsvergleiche/desi.html).

such as guidelines and criteria for assessment, the functions of assessment in RFL classes, the frequency of assessing different skills and the frequency of using different task types and test formats. Part 3 consisted of three questions concerning RFL teachers' wishes and needs regarding the in-service training of assessment procedures. The questionnaire can be downloaded from the IRIS database (https://www.iris-database.org).

Ten out of 28 questions were in the open-ended format, which allowed teachers to introduce new content and express their opinions beyond the answer options provided. The answers to the open-ended questions were put into different groups by each individual author, then a common categorization was agreed upon, thus the nomination of the categories derives from the authors. For example, answers such as communication, oral communication, and communicative skills were combined into a single category labelled "communicative competence".

The survey was developed in esurv (https://esurv.org) and distributed primarily through national and regional associations of RFL teachers in the four countries. The data collection took place in fall and winter 2016/17.

Results

In order to determine the **teachers' workload related to assessment,** they were asked how much time in total they spend – both at school and at home – preparing assessment instruments, developing tests, correcting and marking tests and discussing test performances and test scores (Q16). The results showed that around one third of the teachers spend 20–30% of their time on assessment procedures, and around 20% of the participants each spend either 10–20% or 30–40% of their time on assessment. Teachers who devote between 5% to 10% or more than 50% of their time to assessment procedures make up 10% of respondents in each group.

Regarding their **training in language assessment** (Q11), the results of the survey showed that around 62% of the teachers attended courses at university or during their in-service training that taught them aspects of language assessment, whereas around 38% received no such training at all. It should also be noted that the percentage of teachers who received this kind of training was the highest in the group of teachers who started to teach Russian less than 5 years ago as compared to any other group. This indicates that the topic of assessment has recently started to play a greater role in teacher training than it did in the past. As far as the contents of the assessment courses are concerned, most of the courses dealt with linguistic aspects such as analyzing morpho-syntactic and lexical mistakes and correcting and marking texts.

In secondary school contexts there are certain **legal regulations that determine different aspects of assessment** in RFL classes; therefore, we asked teachers whether they were aware of such

regulations in the specific educational context they work in (Q20). While 51.1% of the informants mentioned that they were aware of the fact that such legal regulations existed, 22.6% stated that there were no guidelines whatsoever that influence their assessment practices. Transferring regulations from other foreign language subjects or from the school-leaving exams was mentioned by about 9% of RFL teachers.

Since assessment procedures should reflect the **learning objectives in the RFL classroom**, the respondents were asked to identify the main goals of their Russian classes (Q13). The two objectives most frequently mentioned by RFL teachers were communicative competence (17.6%) and a basic knowledge of grammar and vocabulary (15.9%). Other frequently mentioned goals were different aspects of multilingualism (12.3%), skills according to CEFR (11.6%), cultural knowledge (9.7%), making learning Russian pleasurable (8.5%), positive attitudes towards and an interest in the Russian language and culture (7.2%), methodological competence, i.e. learning how to learn a foreign language (5.3%), and the ability to perform real-life tasks (5.1%). When the answers 'communicative competence', 'ability to perform real-life tasks' and 'skills according to CEFR' are combined, they amount to 34.1% of all aims listed by RFL teachers, which is twice as much as 'basic knowledge of grammar and vocabulary'.

The results of the survey also indicated that the **frequency of assessment of different skills** roughly corresponds to the hierarchy of teaching objectives (Q24). The respondents had to rank how often they assess different skills on a scale from 1 ('never') to 4 ('often'). Out of the 13 (sub)skills given, RFL teachers reported testing speaking (3.69), writing (3.59), vocabulary (3.57), and grammar (3.50) most frequently. Mediation (2.63), intercultural competence (2.58), audio-visual comprehension (2.49), and sociopragmatic competences (2.25) were ranked the lowest. Among the traditional four skills, listening comprehension is tested least frequently. Yet the reported frequency of assessing different language skills is not reflected in the **use of different task types and test formats** as reported in the answers to Q24. Out of the five test formats that are used by more than 80% of the teachers (answering questions about a text (audio, audiovisual) using keywords, multiple choice, matching, role play, vocabulary tests), only one format can be used to assess speaking, and not a single task type assesses writing. The task types used least frequently, namely by fewer than a third of the teachers, are portfolios, editing tasks and descriptions of pictures in stories.

Since teachers' **subjective theories about a 'good' test** influence their use of tests, their development and correction, we asked RFL teachers to describe what a 'good' test means to them (Q17). A test that corresponds to the instruction and the competence level of the group was the most frequent answer given by the teachers (19.3%). According to the informants, a 'good' test should also target different skills (11.6%), differentiate between weak and strong pupils (11.2%), be transparent

in terms of assessment criteria, marking and task types (10%), include different types of tasks (9.6%), be valid (9.2%) and feasible (9.2%), should provide feedback to teaching (7.6%), and be interesting to students (7.6%). Only few answers referred to useful feedback for learners (2.8%) as a 'good' test characteristic.

Only if teachers are aware why they actually assess their learners and how to use the information gained in this way, can assessment be valid and useful. As shown in Figure 1, RFL teachers mention 13 different **purposes of assessment**, most of which are also mentioned in the research literature on assessment. The majority of the answers given to Q15 refers to assessment of learning, in other words, summative*e assessment (19%), and using assessment as feedback to inform instruction (16.7%). The category ranking third is 'assessment of knowledge' (15.8%), comprising answers such as 'testing of lexis and grammar' without explicitly mentioning a definite purpose.
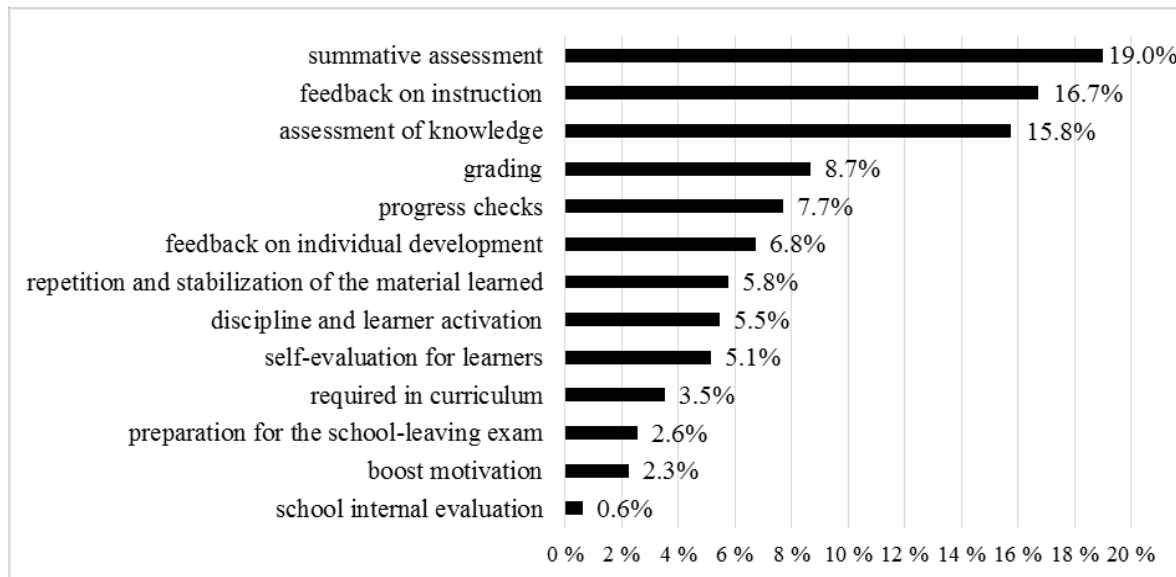
Figure 1. Functions of assessment.

A large percentage of the answers deals with assessment *for* learning (28%): progress checks (7.7%), feedback on individual development (6.8%), repetition and stabilization of the material taught (5.8%), self-evaluation for learners (5.1%) or preparing for the final school-leaving exam (2.6%). These results correspond with the teachers' answers to Q21, in which they had to rate **the functions of written and oral exams** according to their importance on a 6-point Lickert scale from 1 (='not important at all') to 6 (='very important'). The four functions that were rated highest deal with both summative and formative assessment: RFL teachers use assessment so that students see what they can do and where they still have problems (M=5.31), to give students individual feedback on their strengths and weaknesses (M=5.26), to check whether students have achieved the learning objectives (M=5.11), and to let them reflect on their own learning process and progress (M=5.09).

When asked **to evaluate different components of language assessment literacy** which are important for their job as a RFL teacher on a scale from 'not important at all' (=1) to 'very important' (=6) (Q23), respondents rated five components highest: making assessment transparent and comprehensible (*M*=5.6), assessing learners in a fair way (*M*=5.4), giving students feedback on the basis of test scores (*M*=5.3), evaluating the impact of instruction (*M*=5.2), and learning about students' strengths and weaknesses (*M*=5.2).

With regard to the **development of tests** (Q25), most of the respondents reported that they create tests and test tasks on their own (25.9% of the answers), adapt tasks from textbooks or the teacher's handbook (22%) or recycle old test tasks (16.4%). In only about 12% of all cases do teachers develop tasks together with colleagues or adapt tasks from the Internet.

When asked to express their **wishes and needs for professional training in the area of language assessment** in an open format (Q26), teachers see the biggest need in the training of developing and assessing four-skill-tasks (25.6%). Various teaching-related topics not connected with assessment (17.6%) come second. The other four most frequently mentioned needs deal with creating new test formats for the school-leaving exam (9.7%) and the training of assessing speaking (9.3%), writing (7.5%), and mediation tasks (7.5%). These results correspond to the teachers' answers to Q27, in which they had to rate a list of topics for professional training sessions on a scale from 'not important at all' (=1) to 'very important' (=6). As can be seen in Table 1, the most important areas where teachers feel the need for further training are the development and evaluation of classroom tests (*M*=4.91), the development and evaluation of school-leaving exams (*M*=4.81), and the use of evaluation grids for assessing speaking, writing and mediation tasks (*M*=4.80). The three areas of assessment in which RFL teachers consider training least necessary are the Russian certificates 'Тестирование по русскому языку как иностранному' (ТРКИ) (*M*=3.61), the use of self-assessment (*M*=3.58), and the use of peer-assessment (*M*=3.50).

Table 1. Topics mentioned for training sessions

Topics	M
Development and evaluation of classroom tests	4.91
Development and evaluation of school-leaving exams	4.81
Assessment scales for speaking, writing and mediation	4.80
Adapting existing tests for one's own classes	4.42
Effective feedback on the basis of test results	4.42
Analysis and evaluation of the conducted tests	4.39
Advantages and disadvantages of different task types	4.34
Testing integrated skills	4.27
Development of test specifications	4.13
Testing intercultural competence	4.11

Evaluation of project work in RFL classes	4.09
Continuous formative assessment in RFL classes	3.86
Russian certificates ТРКИ	3.61
Self-evaluation	3.58
Peer-evaluation	3.50

Discussion of results

Before discussing some of the results in more detail, we would like to mention that we are aware of the limitation of our study, which has so far been conducted by using a questionnaire only. Qualitative analyses such as teacher interviews and classroom observation will undoubtedly have to follow in order to back up our data and reduce the risk of falling prey to social desirability. What came as a surprise in our study is that communicative competence and task-based learning have finally entered the Russian language classroom; this does not correspond with the feedback both authors have received in numerous training sessions from RFL teachers over the years: Russian has to be taught differently because of its morphosyntactic and lexical complexity.

Teachers in Austria, Germany, Switzerland, and South Tyrol are both instructors as well as test developers and assessors. Around 25% of their time is spent testing and assessing student performances, but 38% of the respondents state that they underwent no training in testing and assessing during their university career and/or their in-service teacher training. In our opinion, this calls for an implementation of testing and assessing modules in foreign language teaching curricula. Assessment literacy has only played a minor role in 'Education Land' so far (DeLuca & Klinger 2010; Popham 2011: 267). The fact that nearly a quarter of all RFL teachers think that there are no existing guidelines that might influence their assessment practices suggests that teachers need to be made aware of the relevant regulations and decrees during their studies or in-service training sessions.

It is due to the teachers' double-role as stakeholders that they should equally be taught and trained in the fields of assessment *of*, *for* and *as* learning. The answers given by the teachers in the survey differ according to the type of question format: in open questions they stress the necessity of summative assessment and its importance as feedback on their teaching, while in closed questions RFL teachers favour formative and diagnostic assessment and individual learner observation as key purposes of a test. Diagnostic testing as a topic for further training is, however, just an 'also-ran': only about 1% of the respondents mention it. Equally remarkable is the fact that only 2.8% of the teachers consider a test to be a crucial means of providing feedback to the learner.

Since classroom-based assessment is mainly brought about by justification processes in connection with achievement tests (Bachman & Damböck 2017), the role of norm-oriented marking on the one hand and criterion-referenced grading on the other hand needs to be considered and discussed with more weight in educational curricula. In the survey, nearly 20% of the RFL teachers

emphasize that a 'good' test is tailored to the learner group's level of competence and based on the instruction which has preceded it. Despite the fact that communicative competence has finally entered the Russian language classroom, teachers nevertheless feel that a sound knowledge of grammar is necessary, and that all students – even the weaker ones – should pass the tests and exams given. This situation requires a differentiation of tests, which makes it necessary to introduce dynamic assessment forms and focus on the interplay between achievement and proficiency, curricular validity and construct validity in classroom assessment.

In the eyes of RFL teachers, only 9.2% consider construct validity as a characteristic of a good test. According to them, construct validity means testing the learners' 'real competence'. Weak students in one group might be average students in another group – so much for norm-orientation. What really matters is the competence level they should reach according to the CEFR. But if teachers match their tests to their students' individual abilities, they may want to consider whether the decisions of what their learners' 'real competence' is can be consistent and fair. They will also have to consider whether their scores are comparable to the scores of teachers in other schools, especially when they are the only RFL teachers at their schools. According to our survey, on a scale from 1 to 6, highly ranked components of language assessment literacy are transparency and comprehensibility (M=5.6) as well as fairness (M=5.4); therefore teachers need to learn how to set up assessment blueprints, write understandable rubrics and develop valid and objective tasks.

Since RFL teachers may have few opportunities to meet with fellow teachers to discuss relevant language performance features and develop test tasks together at the same school, collaborative work needs to be carried out during professional in-service training sessions. And although RFL teachers state that, on a scale from 1 to 4, they mainly test the productive skills speaking (M=3.69) and writing (M=3.59), which came as a real surprise to the researchers, the test formats they apply most frequently do not back up this information. Only one format named is suitable for assessing speaking tasks: *role-play*. Listening (M=3.17) is the receptive skill that is tested less frequently than reading (M=3.22), which may have to do with the special difficulties this skill entails. This is why teachers need example assessments and knowledge of best practices of how to prepare and develop specific listening test items that fit their purpose, since a 'good' test covers all four skills (as 11.6% of all the answers given suggests).

One of the main differences to Vogt's (2010) and Vogt & Tsagari's (2014) report in our study is that their respondents saw no need for being trained in the area of developing and marking classroom tests. Our respondents, however, on a scale from 1 to 6, acknowledge this area as important (M=4.91) among their wishes and needs for professional training (Drackert & Stadler 2017). As

already mentioned above, the reason for this dissimilarity may lie in the lack of best practice assessment examples and standardized Russian tests calibrated to the CEFR.

References

Bachman L., Damböck B. (2017): Language assessment for classroom teachers. Oxford: Oxford University Press.

Crusan D., Plakans L., Gebril A. (2016): Writing assessment literacy: Surveying second language teachers' knowledge, beliefs, and practices, Assessing writing, 28, 43–56.

DeLuca C., Klinger D.A. (2010): Assessment literacy development: identifying gaps in teacher candidates' learning, Assessment in Education: Principles, policy & practice 17/4, 419–438.

Drackert A., Stadler W. (2017): Beurteilungskompetenz von Russischlehrkräften in Deutschland, Österreich, der Schweiz und Südtirol (DACHS): Zwischen Status Quo und aktuellen Bedürfnissen, Zeitschrift für Fremdsprachenforschung, 1/28, 233–258.

Hasselgreen A., Carlsen C., Helness H. (2004): European survey of language testing and assessment needs. Part one: General findings, EALTA.eu. URL: http://www.ealta.eu.org/resources.htm (retrieved 2017-08-19)

Leung C. (2014): Classroom-based assessment issues for language teacher education, In: Kunnan A. J. (Ed.), vol. 3. The companion to language assessment. Chichester, UK: Wiley Blackwell, 1510–1519.

Mehlhorn G. (2016): Russisch, In: Burwitz-Melzer E., Mehlhorn G., Riemer C., Bausch K., Krumm H. (Hrsg.): Handbuch Fremdsprachenunterricht. 6., Völlig überarbeitete und erweiterte Auflage, Tübingen: A. Francke, 534–539.

Popham W. J. (2011): Assessment literacy overlooked: A teacher educator's confession, The Teacher Educator 46/4, 265–273.

Stadler W., Kremmel B. (2018): Testprinzipien, In: Stadler W., Hinger B. (Hrsg.): Testen und Bewerten fremdsprachlicher Kompetenzen. Eine Einführung, Tübingen: Narr Francke Attempto, 39–55.

Stiggins R. J. (1991): Assessment literacy, Phi Delta Kappan, 72, 534–539.

Tiroler Tageszeitung (2017): Spanisch und Russisch sind im Trend – Englisch ist der Dauer- brenner. Nummer 267-ST, 27.9.2017, 20.

Vogt, K. (2010): Assessment Literacy bei Fremdsprachenlehrkräften – Bedarf und Bedürfnisse in Aus- und Weiterbildung. In: Altmayer, C.; Mehlhorn, G.; Neveling, C., Schramm, K. & Schlüter, N. (Hrsg.): Grenzen überschreiten: sprachlich fachlich – kulturell. Dokumentation zum 23. Kongress für

Fremdsprachendidaktik der Deutschen Gesellschaft für Fremdsprachenforschung (DGFF) Leipzig, 30. September – 3. Oktober 2009. Baltmannsweiler: Schneider, 267–277.

Vogt K., Tsagari D. (2014): Assessment Literacy of Foreign Language Teachers: Findings of a European Survey, Language Assessment Quarterly, 11 (4), 374–402.