TRANSLATION ASSESSMENT: IS THERE ANYTHING TO BE TESTED OBJECTIVELY?

Tetiana Korol
Kyiv National Linguistic University, Kyiv, Ukraine
koroltatyanalq@gmail.com

Due to the complex and integrative nature of translation competence construct, its free of subjectivity assessment is considered quite a challenging task for modern translation pedagogy. The conducted literature review revealed diverse directions and opportunities to raise translation competence assessment objectivity. In this key, there is a need for the further studies focused on the designing and implementing of objective translation tests. Thus, the given paper aims at the development and verification of objective multiple-choice computerised test in Translation for the Financial Sector, which combines key-based grading procedures and involves latent translation phase encouraging students to translate with the use of particular translation transformations engaging proper psychophysiological mechanisms in order to respond the item adequately. This study has a mixed research design and involves the qualitative and quantitative analysis of the developed test performance and comments on it, as well as word dictation, translation declarative knowledge test and subjective translation exam test results of 82 third-year students of Poltava University of Economics and Trade received during 2017–2020 academic years. The findings of this study reported on steady correlation of the results of the developed objective translation test with the subjective exam ones (0.992). This fact gives reason to recommend its application as a supplementary translation competence assessment tool contributing to the assessment objectivity. Further investigation should be directed to the determination of the functions and place of the developed objective test in the integrated system of translation competence assessment.

Keywords: translation competence assessment; subjective testing; objective testing; open-ended test; multiple-choice test; test verification; translation in the financial sector; declarative translation knowledge; procedural translation knowledge.

Introduction

Expanding demand for highly qualified philologists able to provide efficient translation and interpretation services in different spheres boosts the theoretical and practical research in modern translation pedagogy aimed at the prospective translators' university training optimisation. Translation competence assessment righteously appears to be a particular sticking point in this area. On the one hand, it should arrange university training process, monitor and promote translation competence acquisition, while, on the other, assessment procedures applied should provide objective and reliable information on students' translation skills and learning outcomes with high-stake washback effect resulting in their course grades or even Bachelor's degree certification. No wonder, that the problem of the development of such translation assessment tools grows in its topicality.

In response to this trend, both scientists and practitioners search for versatile assessment means to meet lots of different educational and social requirements. The task gets more complicated due to several factors:

1. The lack of generally accepted and recognised translation competence structure and acquisition models that makes assessment construct definition hard. Despite some obvious similarities between leading up-to-date translation competence model suggested by PACTE group (Hurtado, 2015) and EMT competence framework (EMT Competence Framework, 2017) based on thorough big-scale empirical research they still differ in a number of components as well as in their categorisation and interpretation. Moreover, not all of the singled out components of these translation competence models can be easily utilised for domestic translation training and assessment practice, especially in terms of activity-based approach according to which knowledge, sub-skills, skills and sub-competences should be distinguished and attributed to particular training stages (Chernovaty & Kovalchuk 2019).

2. The necessity to gauge translation process, product and service parameters from different perspectives for the sake of obtaining objective, comparable and informative results (Kockaert & Segers, 2017). This fact urges to search and develop varied assessment tools to provide cross-assessment (triangulation) at different stages of the translation training process (Hansen, 2008).With this aim, such supplementary periactional assessment techniques (Krings, 2005) as verbal and visual data elicitation in the process of translation (e.g. individual / pair / team think-aloud protocols, translation diaries, translation surveys and checklists, eye-trekking protocols, etc.) and postactional assessment tools (target text quality analysis and evaluation, students' notes and corrections analysis, target text peer-editing, retrospective verbal protocols (RVPs) and so on) are being suggested, developed and implemented.

3. The current trinary assessment trend (hetero, self- and peer-assessment) also requires the search and development of new consistent practical assessment means appropriate to its aim and direct user;

Korol, T. (2020). Translation assessment: is there anything to be tested objectively? Avanced Education, 15, 32-40. DOI: 10.20535/2410-8286.200720

4. The need for objective and practical assessment tools providing any tester / rater (either a teacher or a student) with fast, reliable and simple grading procedures.

5. The terms and conditions of modern Ukrainian university translation training are rather tough from the viewpoint of in and out of class activities correlation stipulated by the syllabi. This forces teachers to develop and apply different self- and peer-assessment means as well as practical hetero assessment procedures involving relevant information technologies (e.g. open-source distance learning platform Moodle, Google Docs apps, in-class didactic software such as JoyClass with integrated testing and assessment facilities, and specific testing software such as Hot Potatoes, Open Test 2, etc.).

6. The concept of translation training in Ukrainian university. Its Bachelor's degree curriculum provides for the concurrent acquisition of both foreign language and translation competence not only in social and political domain but several specialised ones. This raises the problem of students' insufficient language skill level and lacking background knowledge as essential prerequisites for proper domain-specific translation performance in the training process and requires transdisciplinary approach application to prospective philologists' instruction and assessment (Saienko, Simkova, 2019).

One of the productive ways to overcome abovementioned difficulties is seen in the development and combination of various testing techniques to be used for translation competence assessment at different training stages and environments. All of them are traditionally subdivided into subjective and objective ones. In terms of translation assessment subjective tests seem to be more authentic and real-life, since they require target text production based on the direct activity performance by the students calling for their creativity and problem-solving strategies. Such tests are conventionally associated with high subjectivity. Since the main ways to evaluate the translation product, i.e. produced target text, are seen in the use of either analytical, holistic assessment scales, grids and procedures or even combined ones.

Analytical assessment scales being based on the evaluation of particular target text properties and aspects comprise error analysis methods and techniques as well (Hurtado, 2015). The last ones may be practised manually on paper or even in e-format with the help of different CAT tools providing students with the efficient feedback (Yang et al., 2017). Anyway, they are not considered to be practical and workable enough but rather time-consuming and still subjective mainly due to the observed disagreement of error classification and weighing (Eyckmans et al., 2009). Holistic assessment method applies descriptors for different either translation competence or performance levels (Waddington, 2001). Its drawbacks include the lack of intra- and inter-rater reliability, assessment lenience and ignoring mistakes and errors (Akbari & Shahnazari, 2019). Combined assessment method tends to reduce assessment subjectivity penalising students for the errors and mistakes made in the process of translation and determining rater's general impression of the received target text (Waddington, 2001). It should be noted here, that in any case, teachers, saying nothing of the students, run a risk of producing subjective inconsistent results and grades, being involved into energy- and time-consuming assessment procedures. The efforts to increase the objectivity of the production translation tests resulted in the recent development of Calibration of Dichotomous Items (CDI) method (Eyckmans et al., 2009), Preselected Items Evaluation (PIE) method (Kockaert & Segers, 2017; Eyckmans, & Anckaert, 2017) and Calibrated Parsing Items Evaluation (CPIE) method (Akbari & Shahnazari, 2019), engaging complicated statistical analysis procedures not appropriate for routine in-class assessment yet.

Speaking about objective testing techniques for translation assessment multiple-choice format takes the leading position here. Basically, such tests require to choose a proper translation for the source text passage presented in the test task stem from the versions provided. They do not involve any production, only evaluation, recognition and selection of the correct answer. It is obvious that such activity has very little in common with real-life translation situation, while translation competence may be assessed and measured in and through performance only (Abdellah, 2007). In addition, in such testing situation, the testees' creativity is not addressed at all (Golovar, 2012). In our opinion, it can be useful and justified for the assessment of students' discrete reviewing and editing skills only.

However, several attempts were taken to study the opportunity to replace subjective translation testing with an objective one. Among them is the empirical research conducted in Iranian university aimed to examine the correlation of the results received from Master's degree candidates' performance in multiple-choice translation test and open-ended one. Multiple-choice translation test dealt with the selection of proper Persian equivalents for the English source text units among the given options, while open-ended test required full translation performance of the same source text into Persian with the use of paper bilateral dictionaries. Statistical analysis of the obtained data showed no significant relationship between the candidates' scores on the multiple-choice and open-ended translation tests and allowed the researcher to conclude that the ability to translate and choose the suggested translation versions do not correlate (Ahmadi, 2011). Similar results were received a year later by Golovar (2012) for Iranian prospective teachers taking translation course.

Kuhn (2012) compared students' performance on open-ended translation test and the test in detecting and correcting mistakes in relevant target text involving different language pairs. The last one was key-based and concurrently checked up automatically by computer software and by raters. The received results did not reveal a statistically significant correlation between high performance in detecting and correcting mistakes in the target text and translation test.

As the considered examples go there is no real opportunity to replace subjective translation performance test with completely objective key-based one. Since these different test types measure different constructs. In other words, despite its high subjectivity the first one at least provides information on the particular text translation performance in the given testing environment. While the key-based test usually focuses on separate aspects of translation performance and even the sum of such test tasks cannot result in the reliable and valid assessment of the candidate's real ability to translate.

The idea of using objective test tasks as a supplementary tool to reduce the subjectivity of translation-based tasks seems to be obvious enough taking into account all mentioned above. It can be followed in the research dedicated to the development of certification examination for court interpreters, where activity-based tasks in different types of interpretation were graded on the basis of holistic scales and accompanied with the multiple-choice test in detecting and correcting mistakes in the target text patterns. In this case, complex grade appeared to be more grounded and objective according to the statistical analysis results (Muratova, 2006).

A similar approach is implemented in SEVTE test for FBI officers, where multiple-choice tests in choosing proper translation of a single word or phrase presented in the context and error detection in the items written in the target language only, accompany varied production translation tests (Abdellah, 2007).

So the efficient and informative translation assessment should involve not only activity-based or production translation assessment tests but also objective key-based ones. The last will prevent too high assessment subjectivity, help monitor translation course content acquisition by the students (Ma, 2014) and provide essential information on the translation process.

That is why **the aim of the given paper** is to develop and verify a computerised objective multiple-choice translation test in the financial domain as a supplementary tool for traditional subjective translation / interpretation tests traditionally used at the course-end examination.

In this context we would like to answer the following research questions:

1. In what way is it possible to develop multiple-choice translation test whose completion involves students' translation performance and reflects its peculiarities influencing translation product quality?

2. Is there any correlation between students' performance on progress word dictations and declarative translation knowledge tests, developed summative multiple-choice translation and open-ended translation or interpretation tests?

Methods

Research Design

In this survey, a mixed research design was used. The descriptive qualitative research method allowed to get better insights into controversial nature of translation testing with the help of both subjective and objective translation techniques and their combination. On its basis, the concept of multiple-choice translation test was developed with its further verification conducted with the help of quantitative method. Qualitative research also involved students' interviewing on their attitude to the provided multiple-choice translation test tasks to gain a better understanding of the variables having influenced the received testing results. The quantitative research method was directed to the processing of the students' grades received for the 7 progress word dictations based on the thematic wordlists, their results in the 3 borderline translation declarative knowledge key-based tests, developed multiple-choice translation test and open-ended translation examination tests. This method was also utilised to carry out the statistical analysis of the received data, which was finally interpreted with the help of qualitative research method.

Participants

The research participants were 82 third-year students, mostly females, aged from 18 to 20 years, majoring in Philology, who studied at the Institute of Economics, Management and Information Technologies of Poltava University of Economics and Trade during 2017–2020 academic years and gave their volunteer consent to participate in this study. In the first term of their third bachelor year, they took the course of Translation Practice from English in parallel with the course of Translation in the Financial Sector.

Data Collection and Procedure

In order to verify the developed multiple-choice test we decided to follow the students' progress of the acquisition of the underlying components of bilingual sub-competence and world knowledge and thematic one according to PACTE group model, i.e. the knowledge of English and Ukrainian financial domain-

specific terms, notions and concepts, with the help of traditional Ukrainian to English word dictations held for every thematic unit (7 times during the term) and assessed according to 100% scale. The received grade depended on the correlation of the correctly written vocabulary units and a total number of the words included into the dictation. At the end of the term, each student received an average grade for the series of the word dictations written according to 100% scale. The students who missed at least one dictation were excluded from the research.

Matching and open-gap filling key-based tests aimed at the assessment of declarative translation knowledge acquisition (attributed to knowledge of translation in PACTE group model) were held 3 times during the term and graded on the basis of the percentage of correctly given answers (where the best performance corresponds to 100%). According to the curriculum, they comprised theoretical material on lexical and lexical grammatical translation transformations. At the end of the term, each student got an average grade on the series of declarative translation knowledge tests. Any student who has not passed at least one of the three tests was excluded from the study.

Before the examination students passed the computerised multiple-choice translation test containing 40 items automatically selected from 50 ones and presented with the help of OpenTest 2 software. The item and option presentation order was generated randomly and individually for each student during the testing procedure. Time allotment for the test performance was limited to 40 minutes. Each item had only one correct answer. Students were instructed to skip the items which they were not able to respond during the testing and not to choose the option blindly just for guess. In the first year of the research, all the received answers were analysed in terms of item formulation quality and expected difficulty level, distractors selection and appropriateness and were redesigned and corrected correspondingly. The obtained results were calculated according to 100% scale (40 correctly answered items). After testing, students were interviewed. They were asked about their impressions of the test in general, its difficulty level, problems they faced. The students who spent less than a half of testing time on the test performance were excluded from the research.

Finally, students passed examination in Translation for the Financial Sector. It contained four activity-based translation tasks: 1) equivalent translation of the separate sentences from Ukrainian into English assessed on the basis of the criteria of vocabulary and grammar appropriateness; 2) gist translation of the English financial article into Ukrainian evaluated on the basis of the analytical scale for gist translation assessment developed by Shevelko (2016); 3) sight translation of the English text in the financial domain into Ukrainian; 4) consecutive interpretation of the English report in the financial sector into Ukrainian. The last two activity-based exam tests on sight translation and consecutive interpretation were rated on the basis of holistic assessment scale developed by Waddington (2001) (p. 315). For the sake of convenience exam rates were summarised and presented according to 100% grading scale.

As a result, this procedure allowed to collect four categories of numeric data for each student to be compared:

1) financial word dictations average results;
2) translation theoretical declarative knowledge tests average results;
3) multiple-choice translation test results;
4) total exam translation tests results.

They were compared and processed with the help of software Statistica 10.0 by StatSoft, which provides comprehensive data mining, statistical analysis and visualisation procedures. The received result correlation was analysed with the help of the Pearson Correlation Coefficient (PCC). It is used to discover the strength of a linear correlation between two variables x and y. Its values range from +1 to -1, where 1 is totally positive linear correlation, 0 corresponds to the absence of linear correlation, and -1 is the evidence of totally negative linear correlation.

Results

On the basis of the literature review presented above we decided to supplement our translation- and interpretation-based tasks for the course-end exam in Translation for the Financial Sector with the set of objective multiple-choice translation test tasks and study its correlation with other assessment types. The main ambition here was to develop a test task whose successful completion requires some kind of cognitive problem solution based on the analysis of the implicitly received translation product. According to this idea translation production should serve as an intermediary tool but not a final activity result being part and parcel of the test completion with other cognitive processes based on it. This idea stemmed from the concept of translation competence acquisition indicators suggested by Orozco & Hurtado (2002) and specific measurement tools developed for their assessment:

1) students' behaviour facing a translation problem (the problem should be identified, relevant translation strategy applied and proper translation solution produced in case of successful translation performance);

2) behaviour related to the translation errors and their detection / correction;

3) notions about translation, i.e. relevant translation declarative knowledge (p. 380).

Taking into account the peculiarities of domain-specific translation instruction in Ukrainian universities we complemented these measurement directions with the assessment of domain-specific vocabulary acquisition closely connected with relevant background knowledge as the essential prerequisite for proper specialised translation. On the basis of the translation problem classification developed by Orozco and Hurtado (2002) and English financial domain vocabulary peculiarities (Korol, 2009) we singled out the following typical translation problems to be solved with the help of translation procedural knowledge application in the process of financial translation performance from English into Ukrainian:

1) linguistic problems (e.g. synonymic translation of polysemantic vocabulary units functioning with different meanings in general and specific domains based on the deep contextual analysis);

2) extra-linguistic problems (e.g. culture-specific vocabulary units (financial realia) rendition with the help of appropriate translation techniques and strategies);

3) transfer problems (mostly connected with grammatical transformation application not mastered by the students at this instruction stage in full but still observable and applicable when dealing with attributive, in particular, elliptic structures, denoting different financial terms);

4) pragmatic problems (translation adaptation to the situation and target audience, etc. related to translation service provision sub-competence in PACTE group model).

In this way, we got an opportunity to combine two indicators of translation competence acquisition singled out by Orozco & Hurtado (2002) in one multiple-choice test task creating specific integrative activity based on the translation problem identification, strategy selection and solution due to the application of the relevant procedural translation knowledge and some kind of cognitive activity based on the received translation product.

The developed test tasks were monolingual and based on the authentic English text fragments on financial and banking issues. According to the translation problem to be solved they were subdivided into four groups.

1. Test tasks based on the linguistic problems and aimed at the assessment of polysemantic term recognition and processing in translation.

For example, *Choose the sentence (a-d) where the word 'security' is used to denote such financial tools as equities, bonds, swaps, futures, etc.;*

a) <u>Numbers showed that investment in foreign securities slowed in November.</u>

b) Reiss used his Brooklyn home as security for the loan.

c) Many Koreans like the financial security of working for giant companies.

d) The bank offers its customers flexible borrowing, usually without security.

(Here and further in the text the correct answer is underlined).

The performance of this test task involves rather intralingual paraphrasing and source text comprehension skills than proper transfer ones and strongly depends on the active vocabulary acquisition and ability to analyse the given context. It should be mentioned here that paraphrasing skills are considered to be an important prerequisite for the efficient interpretation performance (Russo, 2014). This type of test tasks showed approximately 0.6 difficulty coefficient on the basis of empirical verification (it means that about 60% of the testees gave correct answer).

2. Test tasks based on the extra-linguistic problem and aimed at the assessment of the skill to apply relevant translation strategies and techniques for English financial realia and other equivalent-lacking units rendition into Ukrainian.

For instance, *Which sentence (a-d) can be successfully translated into Ukrainian without descriptive translation application*:

a) <u>For example, people with higher incomes pay more in personal income taxes, even though they are less likely to need help through government programs.</u>

b) A bull market occurs when investors feel confident about the market.

c) On the other hand, bear market occurs, when investors lose confidence and believe that the share values are going to fall.

d) The ability-to-pay principle states that those with higher incomes pay more taxes than those with lower incomes, regardless of the number of services they use.

The performance of this test task involves rather the ability to identify the stated translation problem and does not guarantee the student's ability to formulate a proper target language segment. However, due to the provision of four options in the form of different full-sense sentences it generates extra stress and load on the testee's attention span and working memory. This type of test tasks showed approximately 0.5 difficulty coefficient on the basis of empirical verification (it means that about 50% of the testees managed to answer it correctly).

3. Test tasks based on the transfer problems and aimed at the assessment of the skills to apply different lexical grammatical transformations for the equivalent and natural formulation of the source text sense with the help of target language means in accordance with its rules.

For example, *Which sentence (a-d) should be translated successfully into Ukrainian with the help of annihilation:*

a) While the 1-year and the 30-year securities are obligations of the US government, the former matures in one year so that there is no uncertainty about the return that will be realized.

b) Such assets are referred to as risk-free or riskless assets.

c) Risk indicates that you cannot be certain about the profit of your investment.

d) However, it helps to remember that without risk, it is impossible to obtain returns that make investments grow.

The performance of this test task involves the ability to identify the formulated translation problem and again does not guarantee the student's ability to produce the proper target language segment. However, due to the provision of four options in the form of different full-sense sentences it generates extra stress and load on the testee's attention span and working memory. This type of the test tasks showed approximately 0.3 difficulty coefficient on the basis of empirical verification (it means that about 30% of the testees only chose the correct option while being tested).

4. Test tasks based on the pragmatic problems and aimed at the assessment of the testees' sub-competence of translation service provision.

For example, *Which type of interpretation will be appropriate in such a situation: The IMF is holding an annual meeting with the participants from more than 140 countries?*

a) simultaneous.

b) consecutive.

c) sight.

d) whispering.

The performance of this test task involves the students' general knowledge about translation and translation services, which is checked up in a particular subject-related situation. It does not engage testees into translation production but requires relevant vocabulary acquisition. This type of test tasks showed 0.7 difficulty coefficient on the basis of empirical verification (it means that about 70% of the students chose correct answers).

In figure 1 you can see the sample screenshot of student's testing results.

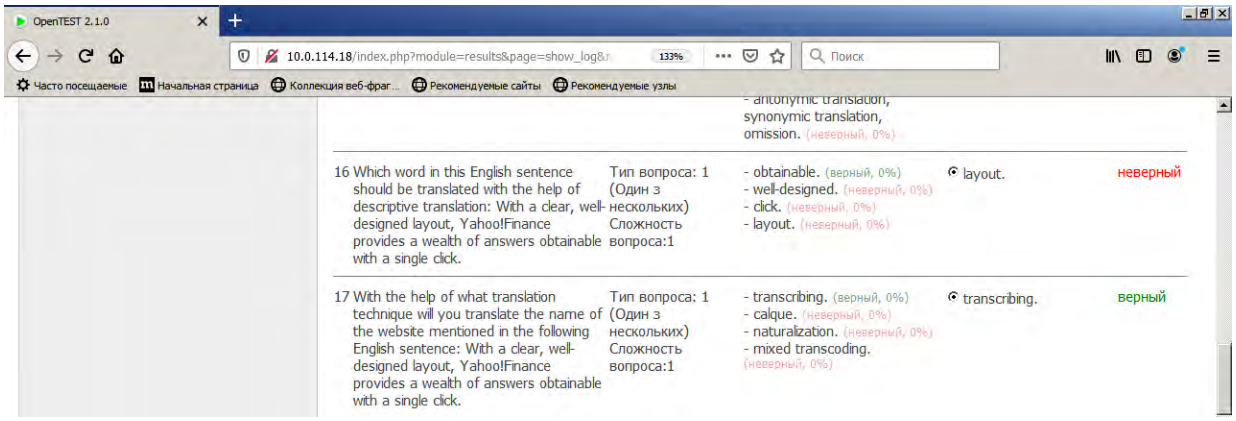

Figure 1. Sample Screenshot of Student's Testing Results in OpenTest 2.

On the test completion, all the students were interviewed. Those who explained their test failures with the lack of financial vocabulary or translation declarative knowledge were excluded from the test result analysis. In general, the presented test tasks were characterised as 'challenging and exciting', some of the respondents named them 'too difficult but still engaging', the others felt 'exhausted' after testing and preferred just perform text translation instead. The listed problems they faced in the process of testing were

connected with the need to read information from the computer screen, to keep in mind several produced translation versions and analyse them simultaneously and to manage the time. Consequently, we may assume that the completion of such test tasks, presented on the computer screen, promoted students' concentration span, developed their short-term and working memory, encouraged to take translation decisions in limited time periods and under pressure. That can be essential from the viewpoint of different types of interpretation performance.

The received results were collected and compared with the students' average word dictation results, average grades on translation declarative knowledge tests and exam performance. The mean values for each type of assessment are summarised in table 1.

Table 1. Mean Values for Each Assessment Type Results

Assessment Type	Number of Students	Mean Value (%)	Minimum Value (%)	Maximum Value (%)
Word dictation	82	77.5	55.0	95.0
Translation declarative knowledge test	82	80.1	59.0	96.0
Objective translation test	82	75.3	53.0	95.0
Subjective translation test	82	79.4	60.0	97.0

As we can see from the table all the assessment types revealed students' sufficient training level > 70 %. This can be explained with the removal of the students who missed word dictations and translation declarative tests as well as those who completed the multiple-choice test in too short period of time and failed it because of the lack of relevant vocabulary and theoretical knowledge from the statistical population being analysed. In this case, objective translation test results appeared to be the lowest (75.3%) which can be explained by the new way of test presentation and procedures of its completion as well as the complex nature of both translation and analytic activities involved being based on the acquired vocabulary and declarative knowledge practical application. The second most challenging assessment type was word dictation with the average value of 77.5%. This can be explained by its directionality from Ukrainian into English and tough spelling and synonymic group provision requirements followed in the assessment procedures. Translation production exam tests showed the average of 79.4% which is explained with its subjective rating conducted by the teacher, who delivered both translation courses to the given students' groups. Finally, translation declarative knowledge test results turned out to be the most successful with the average value of 80.1% since they were directed to the assessment of the particular basic translation knowledge acquisition only.

Finally, the correlation between students' performance on word dictations, declarative translation knowledge tests, developed multiple-choice translation and open-ended translation or interpretation tests was analysed with the help of the Pearson Correlation Coefficient in the software Statistica 10.0 (see Table 2).

Table 2. Correlation of the Results Received with the help of Different Assessment Types

Variable	Correlation of Different Assessment Type Results Marked correlations are significant at p < .05000, N=82					
	Means	Standard Deviation	Word Dictation (%)	Declarative Knowledge Testing (%)	Objective Translation Testing (%)	Subjective Translation Exam (%)
Word Dictation (%)	77.45238	12.74911	1.000000	**0.996580**	**0.989461**	**0.992745**
Declarative Knowledge Testing (%)	80.09524	12.58668	**0.996580**	1.000000	**0.990336**	**0.992511**
Objective Translation Testing (%)	75.29762	13.84752	**0.989461**	**0.990336**	1.000000	**0.992556**
Subjective Translation Exam (%)	79.38095	12.73635	**0.992745**	**0.992511**	**0.992556**	1.000000

So the received Pearson Correlation Coefficient values range from 0.989 to 0.996 approaching +1 for all the assessment results analysed. That is the evidence of positive linear correlation existence between

students' word dictation results, translation declarative knowledge tests, developed objective translation test and subjective translation tests held at the exam. The high correlation of word dictation results and declarative translation testing (0.996) observed can be explained by their similar assessment status and washback effect (progress test), their rather high frequency, reproduction nature of the assignments and students' attitude to them closely reflecting their individual learning styles (see Synekop, 2018). Anyway, it should be stressed here that subjective translation testing carried out at the examination demonstrated an approximately equal strong correlation with other assessment types (0.992). This fact gives reason to recommend their consecutive application as supplementary translation competence assessment tools at different stages of the training process.

Discussion

The received statistical data confirm the idea of the specialised translation quality dependence on the acquisition of domain-specific vocabulary and relevant translation knowledge applicable for the aroused translation problem solution and indirectly comply with the indicators of translation competence acquisition singled out by Orozco & Hurtado (2002). At the same time, domain-specific vocabulary and declarative translation knowledge acquisition are not enough to perform translation properly. These constructs are necessary but not sufficient.

The developed multiple-choice test is based on the creation of a particular testing situation where the testee is required to combine his / her acquired vocabulary, translation knowledge, cognitive and creative abilities in order to solve the formulated problem with the help of translation or its fragment performance as an intermediary step. Translation, in this case, is not presented in the form a product to be assessed, it is latent and implicit. However, since it serves the basis for other decisions and problem-solutions, its potential quality achievement may be deduced on the received results. Such kind of integrative tests should be developed on the basis of the text segments saturated with typical for the particular domain translation problems. From own experience must admit that it is an effort and time-consuming undertaking. The selection of translation problem-containing text segments in order to assess students' behaviour is used in translation problem instruments developed by Orozco & Hurtado (2002) where the testees are supposed to translate the text passage containing those problems and reflect on them responding the relevant questionnaire. In this case, the rater gets an opportunity to find the reasons for the efficient or failed translation solution. This idea is also employed in Preselected Items Evaluation (PIE) method (Kockaert & Segers, 2017; Eyckmans, & Anckaert, 2017), where the segments containing potential translation problems are examined from the point of view of their difficulty and their translation patterns are assessed thoroughly by the raters. The main advantage of the developed multiple-choice translation test is seen in its complex nature connected with the solution of a number of different problems at a time resembling real-life translation situation. However, the same property can be treated as its disadvantage as well. Due to the involvement of lots of variables we cannot be sure which one exactly prevented proper task completion.

The fixed steady correlation of the results received from the developed multiple-choice translation test and subjective translation production one does not relate previously obtained research data (Ahmadi, 2011; Golovar, 2012) mainly due to different test content and rubric. At the same time, the fixed correlation is aligned with the increase of assessment objectivity in case of varied multiple-choice tests application as supplementary tools (Muratova, 2006; Abdellah, 2007).

In case of computerised format, it possesses a number of versatile objective test benefits: preventing students' cheating, automatical objective grading, feasibility for self-assessment, time restriction setting, modelling stressful real-life situations of different types of translation and interpretation, sight in particular.

The next step in this direction should be test task division on the basis of the cognitive load involved for their completion, objective factors determining their difficulty, optimal ways for such test tasks development.

Limitations

Our research was concentrated on the development and application of objective multiple-choice translation test as a supplementary tool to rise subjective translation test objectivity. However, our research had certain limitations which may affect the generalisability of its results. Firstly, the suggested test was focused on the translation in financial domain from English into Ukrainian and provided for the involvement of particular translation procedural knowledge. This means that the findings may be generalised with some caution. Secondly, the study was limited only to the third-year students majoring in Philology of Ukrainian establishments of higher education. Hence, similar studies can be carried out in other domains, language combinations and students' groups.

Conclusions

The conducted research proved the possibility of objective multiple-choice translation test development based on students' cognitive and translation problem-solving activity. Its results steadily correlate with the results of subjective translation production tests in case of prior sufficient domain-specific vocabulary and basic translation knowledge acquisition. Developed multiple-choice translation test covers four types of translation problems to be solved with the help of relevant translation strategies and techniques application. It can be efficiently used as a supplementary assessment tool for balancing out the subjectivity of translation product quality assessment procedures. Further investigation should concern the determination of objective translation multiple-choice role and function in the system of prospective philologists' translation competence assessment.

References:

Abdellah, A. S. (2007). *Developing a process-oriented translation test for assessing English - Arabic basic translation skills*. Unpublished PhD thesis, South Valley University, Cairo, Egypt.

Ahmadi, A. (2011). On the Validity of a Multiple-choice Translation Test as a Substitute for an Open-ended Translation Test in the Iranian University Entrance Examination. *Perspectives: Studies in Translatology*, *19*(4), 307-314. https://doi.org/10.1080/0907676X.2011.556254

Akbari, A., & Shahnazari, M. (2019). Calibrated Parsing Items Evaluation: a step towards objectifying the translation assessment. *Language Testing in Asia, 9(1)*. https://doi.org/10.1186/s40468-019-0083-x

Chernovaty, L. & Kovalchuk, N. (2019). Teaching non-fiction translation: the system of exercises and tasks. *Advanced Education*, *12,* 5-11. https://doi.org/10.20535/2410-8286.155535

European Master's in Translation Competence Framework (2017). Retrieved 20 January 2020 from https://ec.europa.eu

Eyckmans, J. & Anckaert, Ph. (2017). Item-based assessment of translation competence: Chimera of objectivity versus prospect of reliable measurement. *Linguistica Antverpiensia, New Series: Themes in Translation Studies, 16,* 40–56. Retrieved on 20 January 2020 from https://lans-tts.uantwerpen.be/index.php/LANS-TTS/article/view/436/395

Eyckmans, J., Anckaert, Ph. & Segers, W. (2009). The perks of norm-referenced translation evaluation. In C. Angelelli & H. Jacobson (Eds.), *Testing and assessment in translation and interpreting* (pp. 73–93). Amsterdam: John Benjamins.

Golovar, E. (2012). Translator's Performance on Translation Production Tests & Translation Multiple-Choice Tests. *Translation Journal*, *16(1)*. Retrieved 25 January 2020 from https://translationjournal.net/journal/59education.htm

Hansen, G. (2008). Triangulierung im Unterricht: Ubersetzungsprozesse, Eigenkorrektur und Fremdkrrektur. *Informatologia*, *41(2),* 105-115. Retrieved 25 January 2020 from http://hrcak.srce.hr/index.php?show=clanak&id_clanak_jezik=40136

Hurtado A. A. (2015). The Acquisition of Translation Competence. Competences, Tasks, and Assessment in Translator Training. *Meta*, *60(2),* 256-280. https://doi.org/10.7202/1032857ar

Kockaert, H. J. & Segers, W. (2017). Evaluation of legal translations: PIE method (Preselected Items Evaluation). *Journal of Specialised Translation, 27,* 148–163. Retrieved on 20 January 2020 from http://www.jostrans.org/issue27/art_kockaert.pdf

Korol, T. H. (2009). Leksichna movna kompetencija studentiv u chitanni anglomovnoji literaturi jak objekt testovogo kontrolju [Students' lexical linguistic competence in reading for specific purposes as a testing object]. *Visnyk KNLU. Series "Pedagogy and Psychology", 15,* 21–31. Retrieved on 17 April 2020 from http://dspace.puet.edu.ua/

Krings, H. P. (2005). Wege ins Labyrinth – Fragestellungen und Methoden der Übersetzungsprozessforschung im Überblick. *Meta*, *50(2),* 342-358. https://doi.org/10.7202/010941ar

Kuhn, A. J. (2012). *A Study in Computerized Translation Testing (CTT) for the Arabic Language*. All Theses and Dissertations. 3108. https://scholarsarchive.byu.edu/etd/3108.

Ma, L. (2014). Objective Testing for College Translation Course and Its Washback Effects. *Proceedings of the 2014 International Conference on Information, Business and Education Technology*. (pp.188-191). https://doi.org/10.2991/icibet-14.2014.51

Muratova, I. (2006). *Content and Format of the Certification Examination for Court Interpreters as a Methodological Problem*. Unpublished PhD thesis, M. V. Lomonosov Moscow State University, Moscow, Russia.

Orozco, M. & Hurtado, A. A. (2002). Measuring Translation Competence Acquisition. *Meta, 47(3),* 375-402. Retrieved on 20 January 2020 from https://www.academia.edu/6011208/Measuring_Translation_Competence_Acquisition

Russo, M. (2014). Testing aptitude for interpreting: The predictive value of oral paraphrasing, with synonyms and coherence as assessment parameters. *Interpreting*, *16(1),* 1–18. https://doi.org/10.1075/intp.16.1.01rus

Saienko, N. & Simkova, I. (2019). The Impact of Transdisciplinarity on Sight Translation Training in Scientific and Technical Domains. *Advanced Education, 11,* 46-54. https://doi.org/10.20535/2410-8286.164697

Shevelko, K. O. (2016). *Metodika navchannja starshoklasnikiv specializovanoji shkoli z pogliblenim vivchennjam inozemnoji movi perekladu anglomovnih publicistichnih tekstiv* [The methodology of teaching abstract and gist translation of publicistic texts in English to senior pupils of secondary schools specialising in foreign languages]. Unpublished candidate dissertation, Kyiv national Linguistic University, Kyiv, Ukraine.

Synekop, O. (2018). Cognitive Aspect of Learning Style in Differentiated ESP Instruction for the Future IT Specialists. *Advanced Education*, *10,* 40–47. https://doi.org/10.20535/2410-8286.151271.

Waddington, C. (2001). Different Methods of Evaluating Student Translations: The Question of Validity. *Meta, 462,* 311-325. https://doi.org/10.7202/004583ar

Yang, J., Ciobanu, D., Reiss, C., Secară, A. (2017). Using Computer Assisted Translation tools' Translation Quality Assessment functionalities to assess students' translations. *The Language Scholar,* 1-17. Retrieved 20 January 2020 from https://languagescholar.leeds.ac.uk/wp-content/uploads/sites/3/2017/05/Issue-1-YangReissSecaraCiobanu.pdf

Received: February 15, 2020
Accepted: April 25, 2020