

Quantitative Equivalence Level In Poetry Translation

Shakespeare's sonnets translated by D. Palamarchuk

Ihor Kulchytskyi
Applied Linguistic Department
Lviv Polytechnic National University
Lviv, Ukraine
bis.kim@gmail.com

Larysa Tsiokh
Applied Linguistic Department
Uzhhorod National University
Lviv, Ukraine
Larysats@ukr.net

Mariia Malaniuk
Foreign Languages and Translation Studies Department
Lviv State University of Life Safety
Lviv, Ukraine
maria.malanyuk@gmail.com

Abstract— The article highlights the problems of literary, especially poetry, translation equivalence. It has been proved that quantitative indicators are able to determine the main vectors of the analysis idiostyle dominants and their correspondences in translation, and this fact gives grounds to talk about the quantitative equivalence level.

Keywords— *equivalence of translation, quantitative equivalence level, contrastive linguistic statistical analysis, lemmatization, homogeneity criterion, quantitative coefficients of lexical level.*

I. SCIENTIFIC PROBLEM FORMULATION AND ITS SIGNIFICANCE

Being central in translation studies, the problem of translational equivalence has been discussed for almost a century. Scientists offer a wide range of definitions of this concept, which differ depending on methodological approaches, genre and style differentiation of source texts, the authors belonging to different translational schools, etc. The problem of translational equivalence of literary works and poetry in particular is especially critical, because the main problem that occurs during poetry translation and its evaluation relates to the relation between the form of the work and its content.

II. ANALYSIS OF RESEARCH ON THE PROBLEM OF TRANSLATIONAL EQUIVALENCE

Numerous scientists have dedicated their researches to the problem of translational equivalence. Among those there are such well-known foreign scholars as E. Nida, M. Baker, C. Gelferson, J. Kasagrande, J. Catford, A. Neubert and also Ukrainian and Soviet researchers like V. Vinogradov, N. Garbovskiy, R. Zorivchak, V. Komissarov, J. Retsker and others. In spite of detailed studies of the phenomenon, it still remains one of the most controversial in the translation theory. Considering the interdisciplinary nature of modern linguistic research, we consider it appropriate to apply methods of quantitative linguistics to the problem of translation analysis,

which would allow revealing quantitative characteristics of original texts and translations, and thus penetrating into certain motivated quantitative parameters of the features of their content organization.

III. THE PURPOSE AND OBJECTIVES OF THE ARTICLE

The purpose of this article is comparative analysis of Shakespeare's original sonnets and their translation by D. Palamarchuk which aims to establish their equivalence at the quantitative level by means of linguistic statistical analysis of the texts lexical level.

IV. PRESENTATION OF THE MAIN MATERIAL AND THE SUBSTANTIATION OF THE RESULTS OF THE RESEARCH

The traditions of achieving literary text equivalence have evolved over the centuries, so the thoughts of researchers regarding the definition of its concept differ significantly. Some of them, M. Brandes and V. Provotorov in particular, argue that the translation must fully preserve the content and form of the original work, in fact equating the equivalence with total identity [1, p. 5]. There is also the concept of "dynamic equivalence", proposed by E. Nida, which aims to overcome the limitations of a purely semantic approach to equivalence. E. Nida claimed that the translation is to create the translation of "the nearest natural equivalent" of the original in the target language. Dynamic equivalence is the quality of translation, in which the semantic content of the original text is transmitted to the target language in such a way that the reaction of the recipient of translation is basically similar to the reaction of the output receptors [2, p. 119].

Originating from the idea to present a comprehensive description of translation activity in one model, V. Komissarov proposed a theory of equivalence levels. It is based on the assumption that equivalence is established between similar levels of content of the original texts and translation. Translational equivalence may be based on the preservation and loss of various elements of the content contained in the

original. Different levels of equivalence largely depend on which part of the content is being translated. There are five of them, and, according to linguistic tradition, they have a hierarchical structure: 1) linguistic signs; 2) statement; 3) message; 4) the description of the situation; 5) the purpose of communication [3, p. 76].

V. Koller distinguishes five invariant components that influence the choice of different types of equivalences: (1) extralinguistic content of the text (denotative equivalence); (2) connotations, including the style of the work, social and regional peculiarities, frequency of their use (connotative equivalence); (3) textual and linguistic norms (textual and normative equivalence); (4) a reader, for which, in fact, translations are made (pragmatic equivalence); (5) formal aesthetic features of the language of translation, language game in particular, metalinguistic aspects, individual stylistic features (formal equivalence – mainly used in Biblical translations and business correspondence) [4, p. 100].

The basis of the translation analysis is to establish the equivalence of the original text and its translation. Therefore V. Koptilov proposes to apply structural analysis: "... the main method of analysis of literary translation to ensure objective approach to it should be a method of overlaying its structure on the structure of the original" [5, p. 197]. In our opinion, the first stage of such analysis can be the comparison of quantitative characteristics of the original text and its translation at the above-mentioned levels.

Statistical indicators alone are obviously unable to determine the aesthetic features of the translation text. V. Perebyinis notes that statistical characteristics reflect only some of the structural features of texts and can only indicate features of an individual or functional style hidden from simple observation [6, p. 152]. However, quantitative data obtained can determine the main vectors of qualitative analysis of the dominance of the original text idiostyle and its correspondences in translation. We believe there is quantitative level of equivalence between the original text and its translation in comparative analysis Shakespeare's sonnets and their Ukrainian-language translations by D. Palamarchuk. In this study, the translation and the original are compared at lexical level by means of contrastive analysis, which is divided into several stages: complete and accurate description of the language systems under study; both descriptions are compared with each other; in the process of comparison contrastive observations are made [7, p. 6].

At the first stage the original texts [8] and their translations [9] have been converted to electronic form and normalized. Then by means of the AntConc program the number of words and word forms in the original and translation is calculated. A separate word is the sequence of letters between two spaces or punctuation marks; therefore an apostrophe and a hyphen are considered a letter. For further analysis pre-processed sonnets were transferred to the MS Excel environment, where each word acquired corresponding part of speech, its lemma and the number of uses for the Ukrainian and English languages respectively. We use traditional parts of speech classification:

for Ukrainian: major: nouns, adjectives, pronouns, verbs (participles, adverbials), adverbs, numbers; functional: preposition, conjunction, particle, interjection;

for English: major: nouns, adjectives, pronouns, verbs, adverbs, numbers; functional: prepositions, conjunctions, articles, interjection.

It should be noted that we share the viewpoint of I. Vykhoanets [10] concerning the status of the adjectival participle as a kind of a verb-derived adjective with its corresponding categories (gender, number, case) and typical syntactic role as a modifier.

While lemmatizing the following principles were considered:

for Ukrainian: nouns are grouped to singular forms in Nominative case; verbs are grouped to infinitive; all adjectives are grouped to singular, masculine forms in Nominative case, including comparative and superlative forms; adverbs are grouped to base forms, including comparative and superlative forms; case forms of pronouns and numbers are reduced according to the type of declension; the phonetic variants of words are reduced to the initial form (the most frequent one) in cases when the alternation of the initial or final letters is associated with the euphony of the language;

for English: plural and possessive forms of nouns are reduced to the base form; all tense forms of verbs, gerund, Participle I and Participle II are reduced to infinitive; degrees of comparison of adjectives are reduced to positive (base) form.

Specifically compiled program was used to calculate absolute frequency of every lemma automatically.

A comparison of a number of characteristics of the original and translation texts is given in Tables 1 and 2.

TABLE I. GENERAL CHARACTERISTICS OF THE ORIGINAL AND TRANSLATION TEXTS

Statistics	Original	Translation
Text size, (N)	17536	12031
Different word forms, (V_f)	3545	5088
Different words, (V)	2562	3075

The data collected enables to calculate a number of coefficients that may show manifold characteristics to measure the lexical level of the original and translation texts [11; 12; 13; 14].

The richness of the vocabulary, or diversity coefficient (K_d), is the ratio of the vocabulary of lexemes to the text in general (the greater the coefficient is, the more different words are used in a specific text), and it is calculated by the formula:

$$K_d = V/N \quad (1)$$

The average use of a word in the text (K_{wr}) is the ratio of the text to the lexemes and shows how many times each word is used in the text on average and is calculated by the formula:

$$K_{wr} = N/V \quad (2)$$

TABLE II. COMPARISON OF QUANTITATIVE CHARACTERISTICS OF WORD FORMS AND LEMMAS IN THE ORIGINAL AND TRANSLATION TEXTS

Statistical index	Original		Translation	
	word forms	lemmas	word forms	lemmas
Hapax legomena, (V ₁)	2085	1287	3819	1719
Major parts of speech, (PS ₁)	3439	2491	4947	2981
Nouns, (N)	1246	981	1975	1274
Verbs, (Ve)	1139	663	1532	897
Adjectives, (Adj)	731	633	868	481
Adverbs, (Adv)	211	161	257	226
Pronouns, (Pron)	92	38	276	89
Numbers, (Num)	20	15	39	14
Functional words number, (PS ₂)	106	71	141	94
Prepositions, (Prep)	64	43	54	32
Conjunctions, (Conj)	22	14	43	28
Particles (Part)			36	28
Interjections, (Int)	13	11	8	6
Articles, (Art)	7	3		

The coefficient of exclusiveness (hapax legomena) is calculated separately for the vocabulary and for the text; it characterizes the variability of the vocabulary, that is the proportion of the text (vocabulary) occupied by the words that appeared only once:

exclusiveness coefficient for the vocabulary (K_{ev}) is the ratio of lexemes number with frequency 1 (V₁) to the total number of lexemes

$$K_{ev}=V_1/V \quad (3)$$

exclusiveness coefficient for the text (K_{en}) is the ratio of lexemes number with frequency 1 (V₁) to the volume of the text:

$$K_{en}=V_1/N \quad (4)$$

Opposite to the exclusiveness coefficient is that of concentration of the vocabulary (K_{vc}) and the text (K_{nc}), indicating the fraction of the text (vocabulary) occupied by words that appeared 10 or more times:

vocabulary concentration coefficient (K_{vc}) is the ratio of words in the vocabulary with an absolute frequency of 10 or more times (V₁₀) to the total number of words in the vocabulary:

$$K_{vc}=V_{10}/V \quad (5)$$

text concentration coefficient (K_{nc}) is the ratio of words in the text with an absolute frequency of 10 or more times (V₁₀) to the total number of words in the text:

$$K_{nc}=N_{10}/N \quad (6)$$

A relatively small amount of high-frequency vocabulary (low concentration coefficient) and a relatively large number of words with a frequency of 1 (high coefficient of exclusiveness) indicate a significant variety of vocabulary.

Close to the concentration of the vocabulary is the coefficient of lexical density of the text (K_{ld}), which expresses

the ratio of functional parts of speech in the text to the total number of words:

$$K_{ld}=PS_1/N \quad (7)$$

Therefore more lexically dense will be texts with less vocabulary of functional words. It is possible to calculate the coefficients of lexical density both for the major parts of the language as a whole, and separately for nouns, adjectives, verbs, adverbs.

Parts of speech ratios and quantitative indicators of them are considered one of statistical characteristics of literary work lexical level and the parameter of the individual author's style. Those include the coefficient of nominal definition, i.e. coefficient of epithetization (the ratio of noun use to the amount of adjectives), the index of verbal definitions (the ratio of the amount of adjectives to the amount of verbs) [14, p. 128], nominative degree (the ratio of noun use to verb use) [11, p. 50].

S. Zasiiekin points out that psycholinguistic text analysis plays an important role in comparing the original and the translation which can partly be carried out on the basis of our obtained statistical characteristics of Shakespeare's sonnets and their Ukrainian translations by D. Palamarchuk. Thus, the coefficient of verbs (K_{ve}), which is otherwise called the coefficient of aggressiveness, is the ratio of the verbs number and verbal forms (adjectival and adverbial participles) (Ve) to the total number of words, and can be calculated by the formula:

$$K_{ve}=Ve/N \quad (8)$$

The coefficient of embolism (K_{em}), or "litter" of speech, is the ratio of the total number of embols (Em) (words that do not carry semantic meaning), to the total number of words in the text [12, p. 94], and can be calculated by the formula:

$$K_{em}=Em/N \quad (9)$$

The comparison of the above coefficients for the original text and the translation is given in Table 3.

TABLE III. QUANTITATIVE CHARACTERISTICS OF LEXICAL LEVEL OF THE ORIGINAL AND TRANSLATION TEXTS

Coefficient	Original	Translation
Richness of the vocabulary, K _d	0,15	0,26
Vocabulary exclusiveness, K _{ev}	0,50	0,56
Text exclusiveness, K _{en}	0,59	0,75
Vocabulary concentration, K _{vc}	0,09	0,06
Text concentration, K _{nc}	0,65	0,40
Lexical density, K _{ld}	0,20	0,41
Average word repetition in text, K _{wr}	6,84	3,91
Epithetization, K _{nat}	1,70	2,28
Verbal definitions, K _{vat}	0,19	0,17
Nominative degree, K _{nom}	1,09	1,29
Verbal, K _{ve}	0,32	0,30
Embolism, K _{em}	0,03	0,03

To establish significance/insignificance of the statistical difference between the values of the coefficients for the original and the translation, we have calculated χ^2 that is homogeneity criterion in linguistic statistics.

To determine the homogeneity criterion, it is necessary to have a certain number of indicators for each sample.

We use the scheme proposed by V. Perebyinis in [15, p. 72-73] to calculate χ^2 :

$$\chi^2 = S \times \left(\frac{\sum (knTn)^2}{\sum kn \sum Tn} - 1 \right) \quad (10)$$

The indicators for original text (T_1) and target text (T_2) are as follows:

T_1 (0,15; 0,50; 0,59; 0,09; 0,65; 0,20; 6,84; 1,70; 0,19; 1,09; 0,32; 0,03), $\sum T_1=12,36$;

T_2 (0,26; 0,56; 0,75; 0,06; 0,40; 0,41; 3,91; 2,28; 0,17; 1,29; 0,30; 0,03), $\sum T_2= 10,42$;

$S=\sum T=22,76$

Having made the necessary calculations we get:

$$\chi^2 = 0,94$$

To decide whether χ^2 indicates a significant difference, it is necessary to evaluate it according to the table of χ^2 critical values [15, p. 155]. This is done by determining the number of degrees of freedom, which in our case is $f = 11$. The difference is considered significant if the calculated value of χ^2 is greater than the table value for the given level of significance. In our case, 0,94 is less than the smallest number in the series. This means that the differences in the statistical indicators of the original and the translation are insignificant and statistically valid.

V. CONCLUSIONS AND PERSPECTIVES OF FURTHER RESEARCH.

We have defined a number of coefficients characterized the lexical level of the source and target texts. They are: vocabulary richness, exclusiveness coefficient for the vocabulary and the text, the concentration ratio for the vocabulary and for the text, lexical density ratio, average repeatability of different words in the text, the coefficient of nominal definitions, the coefficient of verbal definitions, the nominative degree, verbal coefficient, embolism coefficient. Calculations of the criterion of homogeneity (χ^2) testify that the differences in the statistical indicators of the source and target texts can be considered as inessential and statistically valid. Therefore we can state the quantitative translational equivalence between the texts of Shakespeare's sonnets and their translations by D. Palamarchuk. The logical continuation

of the initiated study will be quantitative analysis of the translations by O. Tarnavsky, V. Marach, I. Kostetskyi, D. Pavlychko, which will expand our perception of their translation style, as well as help to reveal the specifics of the reception of Shakespeare's works in different historical epochs.

REFERENCES

- [1] M. Brandes, V. Provorotov, Pre-translation text analysis. Moskov, NVI-TEZAURUS, 2001.– 224 p. [in Russian].
- [2] E. A. Nida Toward a science of translating// The problems of translation theory in foreign linguistics / Ed. by V.N. Komissarov. – Moskov, Mysl, 1978. – C. 114-137 [in Russian].
- [3] V.N. Komissarov A word on translation (Essay on the linguistic theory of translation). Moscow: International Relations, 1973. - 215 p. [in Russian]
- [4] Koller W. Einfuhrang in die Ubersetzungswissenschaft. Heidelberg, 1983. – P. 186–191 [in German].
- [5] V.V. Koptilov, Primary source and translation: Reflections and observations. Kyiv, Dnipro, 1972. - 216 p. [in Ukrainian].
- [6] V. S. Perebyinis, M. P. Mravytsia, N. P. Darchuk. Frequency dictionaries and their use, Kyiv, Naukova dumka, 1985. – 204 p. [in Ukrainian].
- [7] Y.O.Zhlyuktenko, Contrastive analysis as a method of language study. In: Essays on Contrastive Linguistics, – Kyiv, Naukova dumka, 1979. – P. 5-11 [in Ukrainian].
- [8] Shakespeare's sonnets [The electron source]. – Mode of access: <http://www.shakespeares-sonnets.com/>
- [9] W. Shakespeare, 1966. Sonnets. Translated by D. Palamarchuk. Kyiv, Dnipro, 1966. – 195 p. [in Ukrainian].
- [10] I. R. Vykhoanets, The partial-language status of the participle // Ukrayinska mova. – 2003. – №2. – P. 54-60 [in Ukrainian]
- [11] I. Kamińska-Szmaj, Części mowy w słowniku i tekście pięciu stylów funkcjonalnych polszczyzny pisanej (na materiale słownika frekwencyjnego) // Biuletyn Polskiego Towarzystwa Językoznawczego. – 1988. XLI. – S. 127–136 [in Polish]
- [12] S. V. Zasyekin, Psycholinguistic universals of literary text translation, Lutsk : Volyn. nats. un-t im. Lesi Ukrayinky, 2012. – 272 p. [in Ukrainian].
- [13] S. Buk, A. Rovenchak, Statistical Parameters of Ivan Franko's Novel Perekhresni stezky (The Cross-Paths) // Quantitative Linguistics.— V. 62: Exact Methods in the Study of Language and Text.— Berlin; New York: de Gruyter, 2006.— P. 39-48.
- [14] M. Ruskowski, Wskaźnik epitetyzacji w badaniach stylistycznych // Respectus Philologicus. – 2004. – № 5(10). – S. 48–53 [in Polish].
- [15] Perebyinis, V. S. Statistical Methods for Linguists. Vinnytsya, Nova knyha, 2002. – 168 p. [in Ukrainian].