

## Is there such a thing as a translator's style?

Mikhail Mikhailov and Miia Villikka  
Department of Translation Studies, University of Tampere, Finland  
{Mihail.Mihailov, Miia.Villikka}@uta.fi.

*Like women, translations should be either beautiful or faithful* (cf. Mounin 1994)

### 1. General

It would be naïve to think of a literary translation as an exact copy of the original, its simple replica in another language. Even a situation in which two translators are given the same source text and instructed to translate it as faithfully to the original as possible would result in two clearly different translations.

In translation studies the question of loyalty towards the original is one of the most discussed ones. It has been widely argued that literary translations should maintain the style and structure, the “spirit” of the original as intact as possible (see e.g. Chesterman 1997, Nida & Taber 1974). After all, it is the original author's name that is printed on the very cover of the translation, i.e. even the translation is considered to be a work of the author, not of the translator.<sup>1</sup>

However, it must be borne in mind that translators are also individuals and it is impossible for them to totally set aside their own personality and “get under the original author's skin”. There are several factors influencing the translation process and, consequently, the final product, the translation.

One important factor – the key focus area of this research - is the fact that many translation solutions must be decided upon independently. Translating is not just a simple decoding-recoding action performed merely on the level of words – and even if it were, even the best of dictionaries could not by far offer equivalents and examples to every possible context of a certain word or expression. Also, there are often more than just one suitable equivalent. Usage of these variants is in most cases solely up to the translator's choice.

Further, relations between two different language (grammar) systems are no more stable. Rather seldom is there only one possible way to translate a certain structure. Naturally, every translator aims at producing as fluent a text in the target language as possible, but even opinions on fluency, let alone the means of achieving it, vary greatly between different individuals.

### 2. “Stylistic fingerprints”

The ‘stylistic fingerprint’ problem is nowadays widely discussed in applied linguistics. Most scholars agree that every author has a unique and identifiable style. However, there is no shared opinion on the criteria which can be used for authorship attribution. For example, proportion of nouns or adjectives, and so called “marker words” (e.g. *while - whilst, upon - on*) are frequently used. New methods like principle components analysis of the most common words, measuring vocabulary richness and even letter frequency analysis are being developed (Holmes & Forsyth 1995; Tweedie & Opas-Hänninen 2000).

However, the existence of translator's stylistic fingerprints is less self-evident. It is inevitable that the translator is always under strong influence of the original s/he is translating. Still, is there something personal the translator adds to the target text? Is it possible to define this *something* and to use it to identify the translator?

The aim of the research carried out at the Department of Translation Studies of the University of Tampere is to find out whether translators also have ‘stylistic fingerprints’. The research is based on a parallel corpus of Russian fiction texts and their translations into Finnish. We compared, on the one hand, original Russian texts written by the same author and by different authors and, on the other

---

<sup>1</sup> Recent theories of translation (and literary theories, as well) have strongly questioned the authority of the original author (e.g. Oittinen 1995). However, the translations analysed in this research have mostly been produced around mid-20<sup>th</sup> century and thus we find it appropriate to look at the translations against the theoretical framework and background of their own time.

hand, analysed Finnish translations of different texts performed by the same translator and, in one case, translations of the same text performed by different translators.

### 3. Vocabulary richness

One of the widely used methods of authorship attribution is vocabulary richness measures. D. Holmes and R. Forsyth (1995) used the following quotients for their analysis of the Federalist Papers:

$$(1) R = \frac{100 \text{Log} N}{1 - \left(\frac{V_1}{V}\right)}$$

$$(2) K = \frac{10^4 \left(\sum_{i=1}^{\infty} i^2 V_i - N\right)}{N^2},$$

$$(3) W = N^{V^{-a}}$$

where N is the text length, V the total number of different words used in the text,  $V_1, V_2, V_i$  the number of words used 1, 2, i times,  $a = 0.172$ .

The higher the number of words which were used only once (hapax legomena), the higher is the R quotient. The more high-frequency words in the text, the higher is the K quotient. The more different words there are in the text, the higher is the W quotient.

We used these quotients in our research. The R, K, and W values were calculated for different original Russian texts and for Finnish translations.

**Table 1. Vocabulary richness. Russian texts<sup>2</sup>**

Title	R	K	W
R1	1107.778	58.414	8.307
R2	1087.733	59.892	8.941
R3	1100.168	60.117	8.183
S1	1026.101	50.405	9.003
S2	1073.588	47.996	8.425
S3	1094.196	47.124	7.834
S4	1023.956	50.285	8.839
S5	1078.889	42.718	8.058
T1	1067.322	46.582	8.191
T2	998.055	49.245	8.638
T3	982.655	43.654	8.835

Table 1 shows that the values of R, K, and W are not identical for the same author (however, in most cases they are quite close, e.g. cf. R1 vs. R2, S1 vs. S4). Still, texts by different authors might sometimes have pretty close values of vocabulary richness measures, e.g. Juri Trifonov's novel *Dom na naberezhnoj* (The building on the Embankment, T1) is closer to Arkadi and Boris Strugatski's *Piknik na obochine* (Roadside Picnic, S2) than to other works by Trifonov (T2, T3). The Strugatskis' works fall into two groups, S1, S4 and S1, S3, S5, which differ notably from one another (one possible explanation might be a different degree of participation of the two co-authors). Only Rasputin's works really demonstrate close vocabulary richness. Thus, it can be stated that the vocabulary richness of a certain author is not a stable factor; variation between early and later works may be explained with the fact that, quite naturally, the author's style changes all the time along with his / her taste, prejudices, habits, and so on.

<sup>2</sup> Texts are referred to with the initial letter of their author's (for translations also the translator's) name (see list of texts below).

The same method was then used to compare translations. Most of the texts were translated by the same person, Esa Adrian; for one of the texts — Dostoyevski’s *Zapiski iz podpolja* (Notes from the Underground) — we have two translations, by E. Adrian and by V. Kallama; and one text — Lermontov’s *Geroj nashego vremeni* (Hero of our time) — was translated by U.-L. Heino.

As demonstrated in Table 2 below, the texts translated by E. Adrian (DA, OA, R1A, SA) have quite different R, K, and W values, which seems to indicate that the vocabulary of a translation is to a large degree dependent on that of the original. This assumption is supported by the observation that the vocabulary richness values of the original R1 (see Table 1) and its translation R1A are quite close to each other (substantial deviation is found in the K-factor only, which can be attributed to the difference between the two language systems). Another important point is that the R, K, and W values for the two Finnish translations of Dostoyevski’s story (Adrian, DA, and Kallama, DK) are almost identical.

**Table 2. Vocabulary richness. Finnish translations from Russian**

Title	R	K	W
DK	1038.76	40.03	8.54
DA	1034.74	40.94	8.48
OA	1021.40	30.17	7.91
R1A	1105.85	44.70	8.09
SA	1092.77	32.70	8.17
LH	1059.37	32.88	8.06

#### 4. Most frequent words

Another method which could help to distinguish different authors could be a study of most frequent words in the texts in question. Our method is based on a comparison of lemmatised word lists (this is because people use lexemes rather than word forms). Two texts are compared by selecting the 40 most frequent words from their word lists. Then the F-index is calculated: 3 points are added for each word with close relative frequencies, 2 points for each word with different relative frequencies, 1 point for each word with quite different frequencies, and 1 point is deduced for each word absent in the other list. The results for the original texts were as follows:

**Table 3. Most frequent words. Russian texts**

Texts	Indexes
R1 — R2	40
R1 — R3	53
R2 — R3	43
R1 — S	34
R2 — S	35
R3 — S	34
O — R1	31
S — O	33

It is obvious that if the texts were written by the same author, the frequent words’ lists overlap and many words have close frequencies. One could assume that the same thing might happen if the topic of the texts is close enough, but it doesn’t. For instance, Rasputin’s novels have the same topic as Shukshin’s short stories (country life); moreover, these authors belong to the same “school” of country prosaists. However, their F-index (35 or 34) doesn’t differ much from that of Shukshin vs. Olesha (33) or Olesha vs. Rasputin (31). The table clearly indicates that if the F-index is less than 40, the texts in question are quite likely written by different authors.

The results of the study of the translations (presented in Table 4 below) were entirely different.

**Table 4. Most frequent words. Finnish translations**

Translations	F-Index
DK — DA	63
SA — R1A	44
LH — DK	32
LH — DA	29
DA — R1A	28
SA — OA	28
LH — OA	26
OA — R1A	25

It is evident that the F-index for the texts translated by the same person is high only if the topic is close enough (Shukshin vs. Rasputin, SA — R1A). In other cases the F-index for texts translated by the same person doesn't significantly differ from those translated by different persons. Confirming obvious expectations, the F-index has its highest value when the two different translations of the same text are compared.

## 5. Favourite words

Every person speaks his/her own idiolect, which means that a certain, more or less unique list of favourite words can be compiled from everybody's vocabulary. Although variation is possible, no dramatic changes can be expected. In this respect, we carried out the following experiment: word lists of the two texts were compared against the data of a large text corpus and two lists of words with frequencies much higher than in the corpus were generated. Then these two lists were compared and the number of coincidences (FW-index) was calculated. The higher this FW-index is, the closer is the language of the texts and the more probable it is that the texts were written by the same author.

**Table 5. Favourite words. Russian texts**

Texts	FW-index
R1 — R2	385
R1 — R3	577
R2 — R3	426
R2 — S	242
R2 — O	148
O — S	124

The comparison of the translations, again, shows that the language of different translations of the same text performed by different people is closer than that of the different translations by the same translator.

**Table 6. Favourite words. Finnish translations**

Translations	FW-Index
DK — DA	360
R1A — SA	74
R1A — OA	71
LH — DA	45
R1A — DA	31
R1A — DK	21

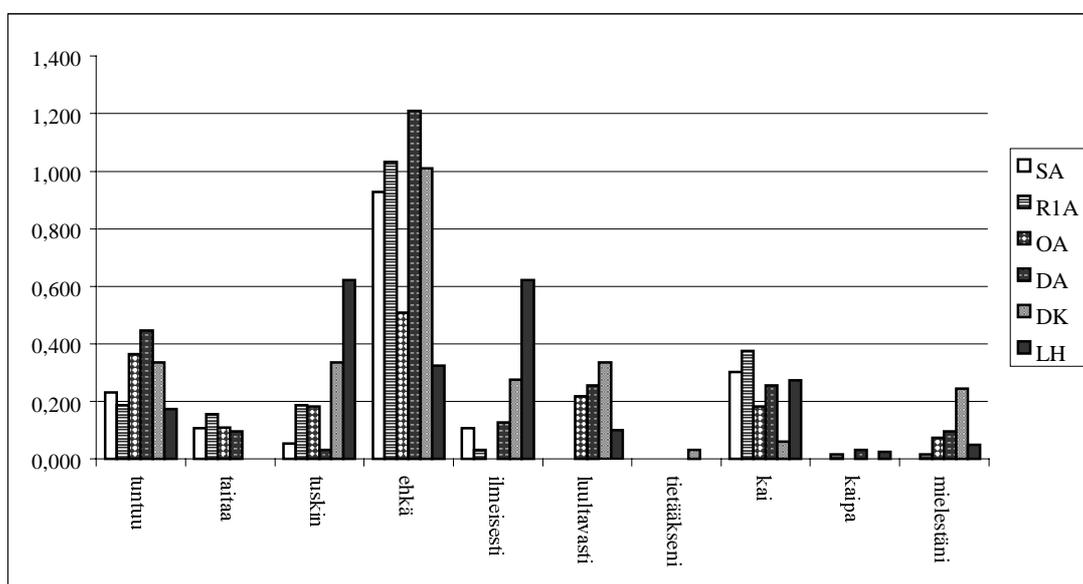
## 6. What is specific?

However, although the translator is a 'chameleon' and the language, style, and core vocabulary of the translation depend on the author's style, we still believe that translator's style is indeed an

existing phenomenon. Despite the strong dependence on the original, all translators have favourite equivalents and patterns of language usage. The analysis of Finnish equivalents for Russian modal markers shows how different translators use different equivalents in analogous situations. Further, the analysis also reveals some patterns of equivalent usage, i.e. certain translators being more fond of certain words than others. This tendency is clearly presented in the different translations of the same text, as well.

This analysis is a good example of the inadequacy of dictionaries as all-embracing guidelines for translators and the inevitability of the translator's own choices. As indicated in Figure 1 below, Finnish equivalents used for the Russian word *kazhetsja* ('it seems (to be)') are *tuntuu*, *taitaa*, *tuskin*, *ehkä*, *ilmeisesti*, *luultavasti*, *tietääkseni*, *kai*, *kaipa*, *mielestäni*. The most widely recognised Russian-Finnish dictionary gives the following equivalents: 1. *näyttää*; 2. *tuntua*; 3. *kai*, *taitaa*. It is worth noting that the first equivalent offered in the dictionary was not used in the translations at all and, on the other hand, the most widely used translation, *ehkä*, is not mentioned in the dictionary at all.

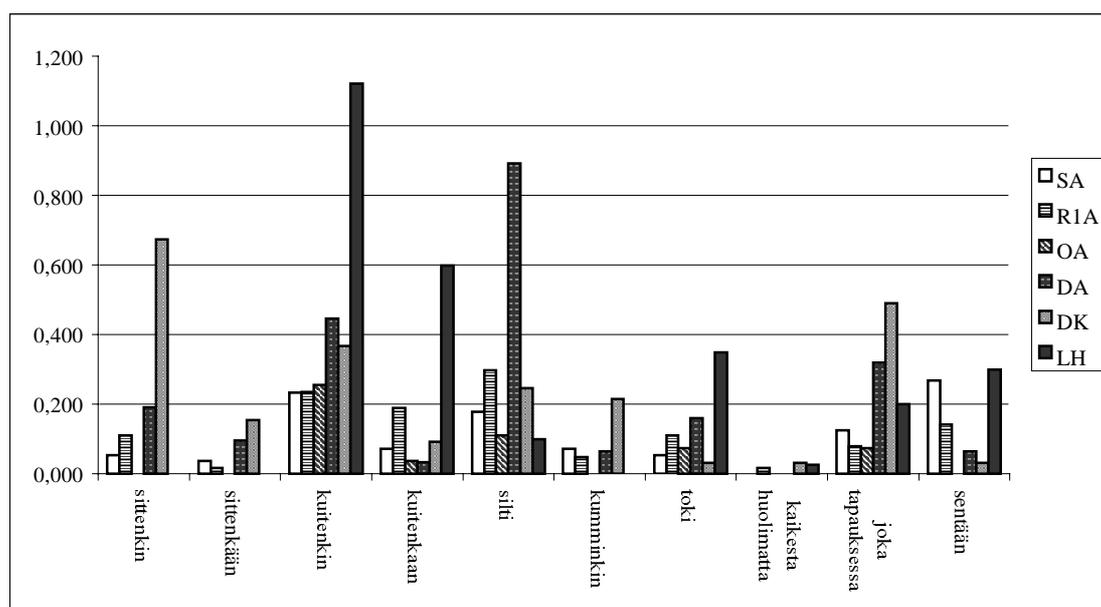
Figure 1. Finnish equivalents for *kazhetsja* in different translations



Based on this data, it can roughly be concluded that translating *kazhetsja* as *taitaa* is typical of E. Adrian; this translation is not used by the other translators at all. Compared to the other translators, V. Kallama seems to prefer the word *mielestäni* and U.-L. Heino appears to be especially fond of the word *ilmeisesti*. An especially noteworthy point is that the frequencies of these words in the two translations of the same novel (see DA and DK) are quite different; thus, their usage doesn't seem to entirely depend on the original.

Similar conclusions can be drawn from the data on the equivalents for *vse-taki* (presented in Figure 2 below). It appears that E. Adrian does not like the expression *kaikesta huolimatta* at all and that *kuitenkin* is mostly typical of U.-L. Heino. It can also be claimed that V. Kallama's repertoire of equivalents is the broadest: he is the only one of these translators who has used all the equivalents analysed.

**Figure 2. Finnish equivalents for *vse-taki* in different Finnish translations**



Interesting results can be discovered by comparing word, sentence, and paragraph counts of originals and translations. The number of words in original / number of words in translation ratio (W-quotient), number of sentences in original / number of sentences in translation ratio (S-quotient) as well as the ratio of number of paragraphs in original / number of paragraphs in translation (P-quotient) are in fact stable values and depend on the pair of languages (Mikhailov 2001). However, table 7 shows that the values of these three quotients are closer for the texts translated by the same person. It is also evident that the two translations of Dostoyevski (DA and DK) differ in this respect. This might be explained by the translator's attitude to the structure of the original: some translators try to generate a text of the same length and the same structure as the source text, some believe that good style and best possible readability in the target language is more important than fidelity to the original.

**Table 7. Words', Sentences', and Paragraphs' ratios for Finnish translations of Russian texts.**

Text	W-quotient	S-quotient	P-quotient
R1A	1,010	0,929	0,975
OA	1,074	0,979	1,001
DA	1,113	0,860	0,975
S4A	1,069	0,953	1,084
S1A	1,073	0,947	1,081
LH	1,033	0,668	0,879
DK	1,069	0,956	0,979

## 7. Conclusions

As was argued in the beginning of this article, it is inevitable that a translator makes a great deal of independent decisions during the translation process. However, when translations were analysed with some widely used authorship attribution methods (e.g. vocabulary richness, frequent words), it appeared as if translators didn't have a language and a style of their own. Still, every translator has a personal set of instruments and stylistic devices. Therefore, in search of the translator's identity (personal features), the most important indicators could be the use of modal words, particles, conjunctions, grammar forms, etc., as well as splitting or joining sentences and paragraphs and expanding or shortening the text.

## List of texts

### Russian original texts<sup>3</sup>

- O: Olesha Ju. *Zavist'* ('Envy')  
R1: Rasputin V. *Zhivi i pomni* (Live and remember)  
R2: Rasputin V. *Poslednij srok* ('The deadline')  
R3: Rasputin V. *Proshchaniye s Matyora* (Farewell to Matyora)  
S: Shukshin V. Short stories.  
S1: Strugatski A. & B. *Paren' iz preispodnej* ('The guy from Hell')  
S2: Strugatski A. & B. *Piknik na obochine* (Roadside Picnic)  
S3: Strugatski A. & B. *Ponedel'nik nachinaetsya v subbotu* ('Monday begins on Saturday')  
S4: Strugatski A. & B. *Popytka k begstvu* (Escape Attempt)  
S5: Strugatski A. & B. *Trudno byt' bogom* (Hard to be a God)  
T1: Trifonov Ju. *Dom na naberezhnoj* ('The Building on the Embankment')  
T2: Trifonov Ju. *Predvaritel'nyye itogi* ('Preliminary Results')  
T3: Trifonov Ju. *Obmen* ('Exchange')

### Finnish translations

- DA: Dostoyevski F. *Zapiski iz podpolja* (Notes from the Underground). Finnish title: *Kirjoituksia kellarista*. Translator: E. Adrian.  
DK: Dostoyevski F. *Zapiski iz podpolja* (Notes from the Underground). Finnish title: *Kellariloukko*. Translator: V. Kallama.  
LH: Lermontov M. *Geroj nashego vremeni* (Hero of our time). Finnish title: *Aikamme sankari*. Translator: U.-L. Heino.  
OA: Olesha Ju. *Zavist'* ('Envy'). Finnish title: *Kateus*. Translator: E. Adrian.  
R1A: Rasputin V. *Zhivi i pomni* (Live and remember). Finnish title: *Elä ja muista*. Translator: E. Adrian.  
SA: Shukshin's short stories translated by E. Adrian.  
S1A: Strugatski A. & B. *Paren' iz preispodnej* ('The guy from Hell'). Finnish title: *Poika helvetistä*. Translator: E. Adrian.  
S4A: Strugatski A. & B. *Popytka k begstvu* (Escape Attempt). Finnish title: *Pakoyritys*. Translator: E. Adrian.

## References

- Baayen R, Tweedie FJ, Neijt A, Krebbers L 2000 Back to the Cave of Shadows: Stylistic Fingerprints in Authorship Attribution. In *The 12th Joint International Conference of the Association for Literary and Linguistic Computing and the Association for Computers and the Humanities*. University of Glasgow, 21-25 July, 2000, pp. 156-158.  
Chesterman A 1997. *Memes of translation: the spread of ideas in translation theory*. Amsterdam Benjamins cop.  
Holmes DI, Forsyth RS 1995 The Federalist revisited: new directions in authorship attribution. *Literary and Linguistic Computing* 10(2): 111-129.  
Mikhailov M 2001. Two Approaches to Automated Text Aligning of Parallel Texts in Fiction. *Across Languages and Cultures*. (forthcoming).  
Mounin G 1994. *Les belles infidèles*. Presses universitaires de Lille.  
Nida E, Taber CR 1974. *The theory and practice of translation*. Leiden United Bible Societies.  
Oittinen R 1995. *Kääntäjän karnevaali*. Tampere, Tampere University Press.  
Tweedie FJ, Opas-Hänninen L 2000 A comparison of methods for the attribution of authorship of popular fiction. In *The 12th Joint International Conference of the Association for Literary and*

---

<sup>3</sup> If no official English translation of the title was found, our own translation is used in inverted commas.

*Linguistic Computing and the Association for Computers and the Humanities.* University of Glasgow, 21-25 July, 2000, pp. 105-107.