

Computational Linguistics Tools in Mapping Emotional Dislocation of Translated Fiction

Ivan Bekhta^a and Nataliia Hrytsiv^b

^a Lviv Franko National University, 1, Universytetska Street, Lviv, 79000, Ukraine

^b Lviv Polytechnic National University, 12, Stepana Bandery Street, 79000, Lviv, Ukraine

Abstract

This research focuses on the interaction between human and computer in processing text of fiction. It is related to operations as applied to overall managing and creating (p)arallel (t)ranslation (c)orpus (PTC) in regard to English as well as Ukrainian language pair. Corpus linguistics comprehensive tools are used both for processing parallel translation corpus results obtained and for the afterwards analysis. Preliminary findings can be exemplary for better understanding of how corpora helps in the scrutiny of the individual style of an author dominants and in what way it can stimulate more qualitative and faithful rendition in translation. The paper provokes a number of issues on how quantitative parameterization can be useful for translation studies analysis. Also, we elucidate the possibilities of NPL manipulation in regard with tagging, hence, researching emotional dislocation as verbalized in fiction.

Keywords ¹

Translation studies, parallel translation corpus, text mark-up, applied linguistics, translation

1. Introduction

If to talk about deepening the profession-oriented skills within the scope of translation, translator training and Translation Studies areas, the advantage of applied linguistics tools in generation of language corpora in electronic format is nowadays a must rather than a choice, especially when it comprises the sheer volume of fiction to be processed. It is a powerful and promising technology to objectively reconstruct arguments in order to amply exemplify its findings and formulate well-grounded hypotheses.

We briefly discuss key characteristics of quantitative comparative analysis, how it can be used in the practice of translation, and what its primary advances are.

The object of study is the parallel translation corpus of Donna Tartt's novel *The Goldfinch* [18] and its Ukrainian variant by Viktor Shovkun [23]. The subject of the study is statistical features of emotional dislocation the English novel "The Goldfinch" in parallel English-Ukrainian corpora and its correlation to translation studies analysis.

Modern practices and techniques of using markup for composing a parallel corpus has proved its efficiency in relation to systematic and regular processing and analysis of texts; it results in generating novel systems and well-elaborated tools for language processing.

Based on the aquired evidence, certain considerations are made regarding the usage of quantitative comparable analysis for further comparison of ST and TT statistics and ratio findings.

COLINS-2021: 5th International Conference on Computational Linguistics and Intelligent Systems, April 22–23, 2021, Kharkiv, Ukraine

EMAIL: ivan.bekhta@gmail.com (I. Bekhta); nataliia.m.hrytsiv@lpnu.ua (N. Hrytsiv)

ORCID: 0000-0002-9848-1505 (I. Bekhta); 0000-0001-6660-7161 (N. Hrytsiv)



© 2021 Copyright for this paper by its authors.

Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

CEUR Workshop Proceedings (CEUR-WS.org)

2. Prerequisites

2.1. Shortenings

ST – (S)ource (T)ext.

TT – (T)arget (T)ext.

PTC – (P)arallel (T)ranslation (C)orpus.

XML – (E)xtensible (M)arkup (L)anguage.

HTML – (H)yper (T)ext (M)arkup (L)anguage.

NLP – (N)atural (L)anguage (P)rocessing.

EDA – (E)motional (D)islocation and (A)lienation (tag used for PTC in a current research project).

2.2. Theoretical and methodological background

A case study of the novel *The Goldfinch* written by the Pulitzer winner novelist Donna Tartt is chosen. The reason is that it has not yet been researched in English-Ukrainian parallel, especially from statistical angle.

In addition, *The Goldfinch* was awarded with Andrew Carnegie Medal for Excellence in Fiction and was film adapted in 2019.

Typologically, the analysed corpus bares the following characteristics; it is annotated; bilingual; illustrative; literary with (a) a certain author (Donna Tartt), (b) a single author type of a text (Donna Tartt's novel *The Goldfinch*), (c) a single author particular text (corpus of Donna Tartt's novel *The Goldfinch*), (d) a distinct genre of text (novels), (e) a separate period (twenty-first century American novel), (f) a peculiar group (Modern American mysterious and *psychological* thriller), (g) a definite theme (emotional register: dislocation, emotions centered around loss, alienation).

Considering the theoretical, practical and applicable issues in regard to author's idiolect and taking into account the overlap with translator's individual style – profoundly discussed and studied from humanitarian perspective, – the particular difficulty of such approach, among other challenges, is to find a golden middle between philological profiling and multifaceted involvement of mathematical linguistics and applied linguistics tools within human language operations.

That such cases are common is of little surprise since there is often a conflicting need to objectively justify the results and support research elaborations with statistics.

In the method and analysis, modern philological researches heavily rely on computer-centered approaches [1, 2, 6, 9, 10, 11, 14, 17, 24].

Thus, translation studies analysis incorporates corpora-oriented and descriptive approaches along with some statistical methods for probability measurement [3, 4, 5, 7, 12, 19, 20, 21, 22]. In addition, one can make use of the undoubted advantages of digital mark up in the parallel: source text and target text.

In the introductory book [4], in a lucid but well-constructed way, the authors collect versatile ideas of different scholars; on this basis, they clearly prove the importance of text and speech corpora, especially in the modern epoch of linguistic studies.

Authors believe corpus linguistics to be among fundamental domains of applied linguistics [4, p. 88 – 89].

Such processing techniques as word listing and concordancing make it possible to have same data viewed and approached from a variety of angles.

It allows stimulating multifaceted analyses and incorporating researchers to constantly rethink the positions of theirs [19].

Language corpora can empower translation studies textual analysis.

As for the advances, in translation studies it should be mentioned that it stems from the fact that they consist of texts in electronic form, and can thus be stored, shared, and processed in ways that enhance their usefulness and comfortable portable processing as compared hard copy corpora [7, p. 69].

Corpus-based translation studies, to recall, has the potential to be a decentering, dynamic force in translation studies on the whole. In this respect, many scholarly publications have already proven its effectiveness and privileges for the sake of their potential benefit.

J. Munday has it that this is a potentially powerful tool to help analyze translation shifts [12, p. 7]. However, little research is available if to mention quantitative analyses in translated text assessment within English-Ukrainian parallel. Therefore, in this article we opt for describing practical application of NPL (Natural Language Processing) methods [13, 15, 16] to fiction translation and its analysis.

3. Pre-computer procedures

3.1. Text mark-up system

We made sure that tags are correct and error-free, since if the tag is missing, the XML file will not work. At this stage, tagging, based on the context, was carried out.

We compared text samples from the source and the target languages.

In focus – word phrases defining emotional dislocation and alienation.

<p> </p> – paragraph mark

<s> </s> – sentence tag

<loss> </loss> – EDA tag for *loss* (of a beloved relative)

<addiction> </addiction> – EDA tag for *addiction* (drug, alcohol, cigarettes)

<anxiety> </anxiety> – EDA tag for *anxiety* (fear of losing the friend)

<punishment> </punishment> – EDA tag for *punishment* (fear of accusation and arrest by police for stealing the painting)

<rambling> </rambling> – EDA tag for *rambling* (orphanage/homelessness)

<guilt> </guilt> – EDA tag for *guilt* (a burden of guilt for stealing the painting and selling out the fake paintings)

<suicide> </suicide> – EDA tag for *suicide* (suicidal thoughts/depression)

<abandonment> </abandonment> – EDA tag for *abandonment* (unrequited love)

<disappointment> </disappointment> – EDA tag for *disappointment* (father's indifference and disappointment in him)

<obsession> </obsession> – EDA tag for *obsession* (*obsession* with the painting and impossibility to watch it constantly)

<failure> </failure> – EDA tag for *failure* (father's unrealized career)

<eda n=1 > </eda n=1 > – verbalized sample of EDA in the text (n=1; n – numbering, 1 – sequence number)

3.2. Sample of the fragments of the ST markup with focus on EDA

Below we provide the screenshots of the marked fragments obtained. To specify, the tagging procedure was conducted for the whole scope of the original and translated documents.

3.2.1. Original fragment 1

<p><eda><loss><s><eda n=1 >Thanks to my fever I had a lot of weird and extremely vivid dreams, <eda>sweats where I thrashed around<eda> hardly knowing if it was day or night, but on the last and worst of these nights I dreamed about my mother: <eda n=1 > a quick, mysterious dream that felt more like a visitation. </s> <s> I was in Hobie's shop – or, more accurately, some haunted dream space staged like a sketchy version of the shop – when she came up suddenly behind me so I saw her reflection in a mirror. </s> <s> <eda n=2 > At the sight of her I was paralyzed with happiness; <eda n=2 > it was her, down to the most minute detail, the very pattern of her freckles, she was smiling at me, more beautiful and yet not older, black hair and funny upward quirk of her mouth, not a dream but a presence that filled the whole room: a force all her own, a living otherness. </s> <s> <eda n=3 > And as much as I wanted to, I knew I couldn't turn around, that to look at her directly was to violate the laws of her world and mine; <eda n=3 > she had come to me the only way she could, and our eyes met in the glass for a long still moment; but just as she seemed about to speak—with what seemed a combination of amusement, affection, exasperation—a vapor rolled between us and I woke up.</s></loss></pain></p> <p><s>ii.</s> </p>

Figure 1: Citation from "The Goldfinch" [18]

3.2.2. Original fragment 2

<p><eda><disappointment><s><eda n=97 > To think of something happening to my mother was specially frightening because my dad was so unreliable. <eda n=97 > </s> <s>Unreliable I guess is the diplomatic way of putting it.</s> <s>Even when he was in a good mood he did things like lose his paycheck and fall asleep with the front door to the apartment open, because he drank.</s> <s>And when he was in a bad mood – which was much of the time – he was red-eyed and clammy-looking, his suit so rumpled it looked like he'd been rolling on the floor in it and an air of unnatural stillness emanating from him as from some pressurized article about to explode.</s> </disappointment></pain></p>

Figure 2: Citation from "The Goldfinch" [18]

3.2.3. Translated fragment 1

<p><pain><loss><s><eda n=1 >Через лихоманку я пережив безліч дивних і надзвичайно яскравих марень, обливаючись потом, не знаючи, котра година, але в останню й найгіршу з тих ночей мені наснилася мати: <eda n=1 > то було швидке й таємниче сновидіння, схоже на несподіваний візит із потойбіччя. </s> <s> Я був у крамниці Гобі чи, точніше кажучи, у якомусь примарному просторі, радше схематично подібному до тієї крамниці, коли вона раптом зупинилася позаду мене так, що я побачив її віддзеркалення. </s> <s> <eda n=2 >Помітивши її, я заціпенів від щастя;<eda n=2 > то була вона, до найдрібніших деталей, з усім своїм ластовинням, мама всміхалася мені, навіть вродливіша, але ще не постаріла, зі своїм чорним волоссям і кумедним вигином рота, не сон, а її присутність заповнювала кімнату: то була її власна сила, жива інакшість. </s> <s> <eda n=3 >І хоч як мені хотілося обернутися, я знав, що подивитися на неї очі в очі було б порушенням законів її світу і мого світу; <eda n=3 > вона прийшла до мене в єдиний спосіб, у який могла прийти, і наші очі зустрілися в дзеркалі на тривалу мить, коли все завмерло; та коли, як мені здалося, вона хотіла заговорити – з виразом обличчя, що, схоже, був сумішшю радощів, прихильності й роздратування, – на нас накотився туман, і я прокинувся.</s></loss></pain></p>

Figure 3: Citation from " Shchyhol" [23]

3.2.4. Translated fragment 2

<p><pain><disappointment> <s><eda n=97 > Думати, що з моєю матір'ю може статися якесь лихо, мені було особливо страшно, адже сподіватися на батька не випадало. <eda n=97 > </s> <s> Не випадало сподіватися — це ще сказано надто дипломатично. </s> <s> Навіть коли він перебував у доброму настрої, він міг учворити якусь несподіванку — загубити чек своєї платні або напитися й заснути, залишивши двері нашого помешкання відчиненими. </s> <s> А коли він був у поганому настрої — тобто майже завжди, — він ходив млявий, із почервонілими очима й у такому пом'ятому піджаку, ніби цілий день качався в ньому по підлозі, і від нього струменів якийсь неприродний спокій, ніби він був предметом, що перебуває під великим тиском і готовий вибухнути. </s> </disappointment></pain></p>

Figure 4: Citation from " Shchyhol" [23]

3.3. Sample of the created corpus of ST with focus on EDA

See below the screenshots of the sample fragments of the created corpus of the ST with focus on EDA

```
<?xml version="1.0" encoding="UTF-8" standalone="yes"?>
<entry xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance">
  <eda>
    <lg>eng</lg>
    <nu>61</nu>
    <sub>abandonment</sub>
```

- Screenshot of the sample fragment 1

```
    <tx>Every moment of their visit had been torture.</tx>
  </eda>
  <eda n>
    <lg>eng</lg>
    <nu>62</nu>
    <sub>abandonment</sub>
```

- Screenshot of the sample fragment 2

```
    <tx>as hard as I'd tried, I hadn't been able to stay away from Pippa. I'd hovered obtrusively and hated myself for it, so painfully excited had I been by her nearness.</tx>
  </eda>
  <eda>
    <lg>eng</lg>
    <nu>63</nu>
    <sub>abandonment</sub>
```

- Screenshot of the sample fragment 3

```
    <tx>The agony of her hand on my sleeve.</tx>
  </eda>
  <eda>
    <lg>eng</lg>
    <nu>63</nu>
    <sub>abandonment</sub>
```

- Screenshot of the sample fragment 4

```
    <tx>Sometimes in bed—adrift in my sighing, opiated, erotic reveries—I carried on long candid conversations with her: we are inseparable, I imagined us saying (comily) to each other, each with a hand on the other's cheek, we can never be apart.</tx>
  </eda>
  <eda>
    <lg>eng</lg>
    <nu>74</nu>
    <sub>abandonment</sub>
```

- Screenshot of the sample 5

```
    <tx>It was hopeless. More than hopeless: humiliating.</tx>
  </eda>
  <eda>
    <lg>eng</lg>
    <nu>75</nu>
    <sub>abandonment</sub>
```

The same results were obtained for all the EDA in the texts (original and translated).

4. Results

4.1. ST EDA in PTC

Having conducted text mark-up (as depicted in subsection above), there was a spreadsheet created that depicted all the EDA samples from the ST in the chosen subtexts.

1	loss	Thanks to my fever I had a lot of weird and extremely vivid dreams, sweats where I thrashed around hardly kn					
2	loss	At the sight of her I was paralyzed with happiness;					
3	loss	And as much as I wanted to, I knew I couldn't turn around, that to look at her directly was to violate the laws c					
4	loss	I felt a petty, irrational pang of disappointment					
5	loss	skull-cracking headache, practically seeing double					
6	disappointment	To think of something happening to my mother was especially frightening because my dad was so unreliable.					
7	disappointment	Apart from this daily awkwardness, I didn't see him much. He didn't eat dinner with us or attend school functio					
8	disappointment	towards the end of the week, a sketchy note from my dad arrived (postmarked Newark, New Jersey) informin					
9	rambling	Even though the apartment was rent stabilized, getting by without my dad's salary was a month-by-month adv					
10	rambling	cold, wet, with a tooth-crunching headache—it struck me that with my dad gone, no one in the world would b					
11	loss	The living room—normally so airy and open, buoyant with my mother's presence —had shrunk to a cold, pale c					
12	loss	I could feel my heart beating,					
13	loss	I missed her so much I wanted to die: a hard, physical longing, like a craving for air underwater.					
14	loss	Lying awake, I tried to recall all my best memories of her —to freeze her in my mind so I wouldn't forget her					
15	loss	drifting uneasily between dreaming and sleep					
16	rambling	My panic must have been written plainly on my face.					
17	rambling	I was a minor, and that my parent or guardian had to be present at an official interview					
18	punishment	I would simply have to admit it and face the consequences, no matter what the consequences were					
19	punishment	At a couple of points, I was on the verge of blurting it out, in my terror.					
20	punishment	It was complicated. Every time I thought of it my stomach squirmed,					
21	punishment	What if someone found the picture? What would happen to me?					
22	punishment	Yet in my homeless limbo, it seemed insane to step up and admit to what I knew a lot of people were going to					
23	punishment	Where the hell was the painting? Though I was nearly blind with panic, I didn't stop but kept going, down the h					
24	loss	My mother's bedroom was dark and cool, and the faint, just-detectable smell of her perfume was almost more					
25	loss	a million other heart-piercing sights,					
26	failure	It's one of those lifelong regrets. I would have loved to do something with my gift but I didn't have the luxury.					
27	failure	Then—after he'd married my mother—it had all fizzled out. He had a long list of reasons why he hadn't broken					
28	punishment	While we were waiting to get our boarding passes I was stiff with fear, fully expecting Security to open my suit					
29	punishment	I felt closed-in and terrified in the bright press of strangers—conspicuous too, as if everyone was staring at me					

Figure 5: The output of EDA in ST

The Figure 5 provides the visual illustration of the corpus obtained from ST data.

4.2. TT EDA in PTC

Following the ST, there was a spreadsheet created that depicted all the EDA samples from the TT in the chosen subtexts.

1	loss	Через лихоманку я пережив безліч дивних і надзвичайно яскравих марень, обливаючись потом, не знаючи, котра годин
2	loss	Помітивши її, я заціпенів від щастя;
3	loss	І хоч як мені хотілося обернутися, я знав, що подивитися на неї очі в очі було б порушенням законів її світу і мого світу;
4	loss	А проте я почував легкий ірраціональний біль розчарування
5	loss	голова в мене розколювалася від болю, речі двоїлися перед очима
6	disappointment	Думати, що з мою матір'ю може статися якесь лихо, мені було особливо страшно, адже сподіватися на батька не випа
7	disappointment	Окрім цих повсякденних незручностей, бачив я його нечасто. Він не обідав із нами й не бував у моїй школі. Він не грав
8	disappointment	наприкінці тижня, від батька надійшло коротке повідомлення (позначене Ньюарком, штат Нью-Джерсі), де він повідомл
9	rambling	Хоч ми сплачували стабільне комірне за наше помешкання, було надзвичайно складно щомісяця обходитися без платні
10	rambling	змерзлий, мокрий, з болем, від якого розколювалася голова, — до мене раптом дійшло, що, позаяк наш батько нас поки
11	loss	Вітальня — за нормальних обставин така свіжа й відкрита, життєрадісна завдяки присутності моєї матері — тепер здавал
12	loss	Я чув, як гупає моє серце,
13	loss	Я так сумував за нею, що мені хотілося померти: то була тяжка фізична туга, схожа на страждання без повітря під водою
14	loss	Не в змозі заснути, я намагався пережити всі свої найкращі спогади про неї — заморозити її у своїй свідомості, щоб ніко
15	loss	важко переходячи від дрімоти до сну,
16	rambling	На мою обличчі було, мабуть, написано, що я перебуваю в стані паніки.
17	rambling	я неповнолітній, і один із моїх батьків або опікунів повинен представляти мене на офіційному інтерв'ю
18	punishment	Тоді мені доведеться просто визнати свою провину й змиритися з наслідками, хоч би якими вони були
19	punishment	Десь разів зо два я мало не зізнався в усьому, опанований страхом.
20	punishment	Проблема була для мене тяжкою. Щоразу, коли я думав про неї, живіт мені судомило,
21	punishment	А що буде, як хтось знайде картину? Що станеться зі мною?
22	punishment	Але в моєму бездомному становищі мені здавалося надто ризикованим розкрити рот і зізнатися в тому, що багато хто р
23	punishment	Де, в біса, картина? Хоч я майже осліп від паніки, та не зупинився, а йшов далі, вниз коридором, на автопілоті до своєї с
24	loss	У спальні матері було темно й прохолодно, і ледь чутні пахощі її парфумів були чимось більшим, ніж я міг витримати.
25	loss	мільйона інших деталей, що розривали мені серце,
26	failure	Це той жаль, який, мабуть, буде вічним. </s> <s> Я хотів би якось застосувати свій дар, але мені не пощастило.
27	failure	Потім, після того як він одружився з мою матір'ю, його ентузіазм пропав. Він називав довгий список причин, чому йому
28	punishment	Поки ми стояли в черзі, щоб одержати посадкові талони, я весь заляк від страху, чекаючи, що працівники служби безпе
29	punishment	я з жахом відчував, як мене затискає в тугому колі незнайомців, — я стояв наче голий, і всі начебто дивилися на мене.

Figure 6: The output of EDA in TT

The Figure 6 provides the visual illustration of the corpus obtained from TT data.

4.3. Description

Next, we created a new spreadsheet that contained calculations of words and word usages in chosen subtext.

Table 1
Qualitative parameters of EDA

Word usage	ST	TT
Total in the text	294619	268409
In chosen EDA subtexts	2248	2031
EDA verbalization	58	75
Implicit EDA	45	59
Explicit EDA	13	16

In total, 294619 words were used in the original of the novel, and 268409 in the translation of the work. EDA sample parallel translation corpus counts 2248 ST words and 2031 TT words.

Explicit EDA and its grammatical variants occur 13 times in ST and 16 times in TT. Implicit EDA are used 45 times in ST and 59 times in TT.

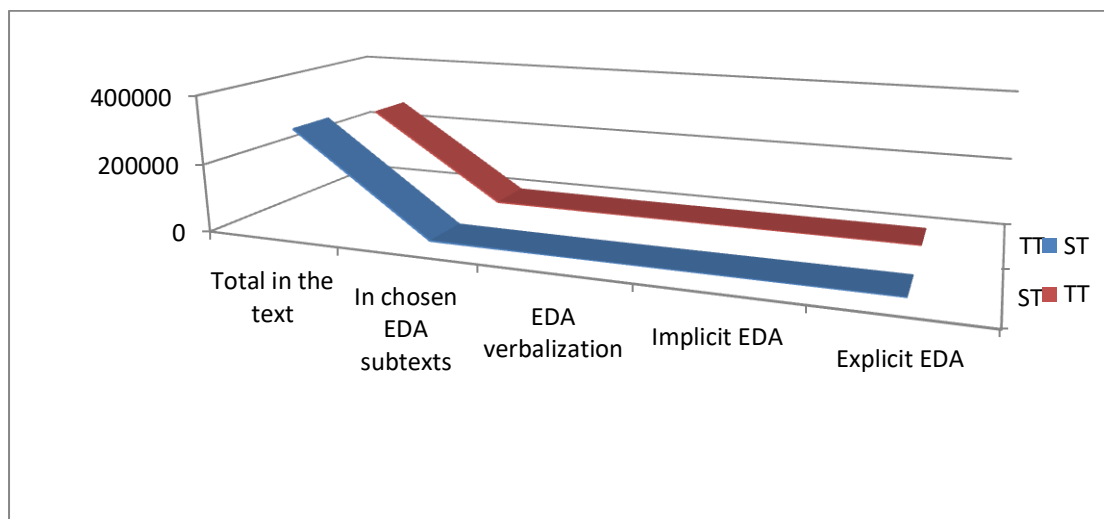


Figure 7: Contrast of EDA in ST and TT

The number of contextually explicit EDA in ST and TT coincides and totals in 107, which comprises a holistic embodiment of the EDA in *The Goldfinch* novel.

Most frequent is EDA verbalized via abandonment and has 23 examples.

The least common is EDA verbalized via failure, it has 2 examples.

Among others there are: loss - 15 examples, addiction - 14 examples, anxiety - 12 examples, punishment - 15 examples, rambling - 7 examples, guilt - 6 examples, suicide - 4 examples, disappointment - 3 examples, obsession - 6 examples.

Below are the percentages of EDA verbalized via abandonment, addiction, anxiety, disappointment, failure, guilt, loss, obsession, punishment, rambling, suicide, when EDA by default remains 100%.

Table 2

EDA percentage rate

EDA	Number (total-107)	Percentage (100%)
abandonment	23	21,5%
addiction	14	13,1%
anxiety	12	11,2%
disappointment	3	2,8%
failure	2	1,9%
guilt	6	5,6%
loss	15	14,0%
obsession	6	5,6%
punishment	15	14,0%
rambling	7	6,5%
suicide	4	3,7%

Now, let us present the visualization of EDA percentage interrelation.

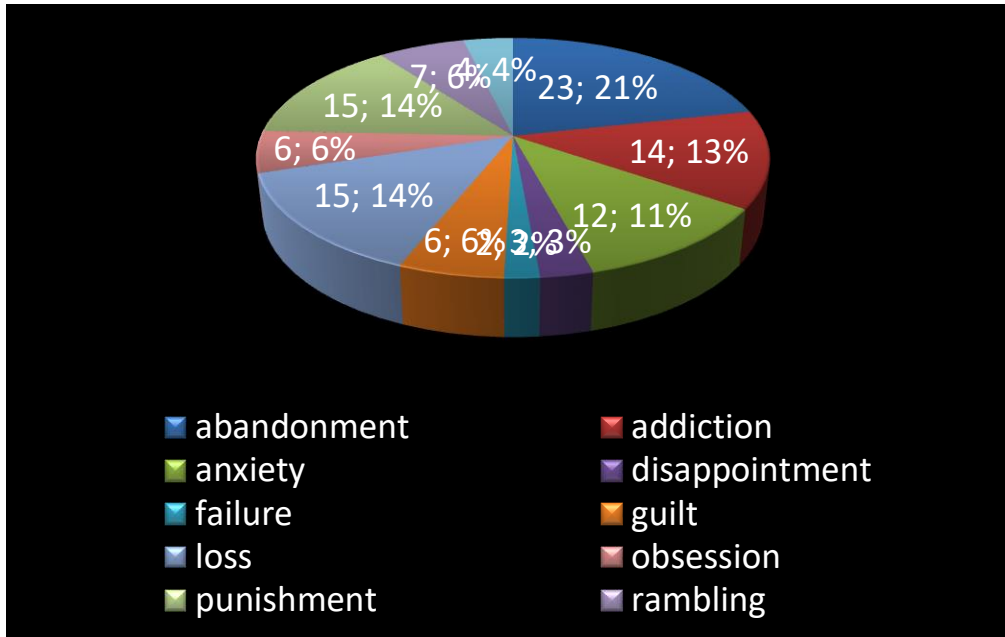


Figure 8: EDA percentage rate

5. Further processing option

5.1. Diversity coefficients

For this data we made use of the available on open access VICTANA platform ([see: http://victana.lviv.ua/nlp/linhvometriia](http://victana.lviv.ua/nlp/linhvometriia)). We focus on contrasting *lexical diversity*; *syntactic complexity*; *speech coherence coefficient*; *exclusiveness index*; *concentration index* as defined in [9]. We are aware that such a comparison of ours is on its embryonic stage if to talk about translational approach; nevertheless, we are convinced that it has the potential to add up to the comparison of statistical profile of the ST and TT, correspondingly. Making use of VICTANA platform we calculate versatile ratio of PTC composed. For this purpose we have taken all 107 EDA samples.

5.2. ST coefficients

N	Coefficient	Input data	Calculations
1.	Coefficient of lexical diversity : $Kl = W / N$	W = 431 N = 886	$Kl = 0.48645598194131$
2.	Coefficient of syntactic complexity : $Ks = 1 - P / W$	P = 29 W = 431	$Ks = 0.93271461716937$
3.	Coefficient of speech coherence : $Kz = (Z + S) / (3 * P)$	Z = 83 S = 69 P = 29	$Kz = 1.7471264367816$
4.	Exclusivity index : $Iwt = W1 / W$	W1 = 330	$Iwt = 0.76566125290023$

		W = 431	
5.	Concentration index : $I_{kt} = W_{10} / W$	W ₁₀ = 12 W = 431	$I_{kt} = 0.02784222737819$

Figure 9: Coefficients of ST

5.3. TT coefficients

N	Coefficient	Input data	Calculation
1.	Coefficient of lexical diversity : $KI = W / N$	W = 303 N = 451	$KI = 0.67184035476718$
2.	Coefficient of syntactic complexity : $Ks = 1 - P / W$	P = 13 W = 303	$Ks = 0.95709570957096$
3.	Coefficient of speech coherence : $Kz = (Z + S) / (3 * P)$	Z = 51 S = 36 P = 13	$Kz = 2.2307692307692$
4.	Exclusivity index : $Iwt = W1 / W$	W1 = 252 W = 303	$Iwt = 0.83168316831683$
5.	Concentration index : $Ikt = W_{10} / W$	W ₁₀ = 3 W = 303	$Ikt = 0.0099009900990099$

Figure 10: Coefficients of TT

Now, let us consider the results in comparison.

Table 3

Coefficients of EDA in Source Text and Target Text

Coefficient	ST	TT
Lexical diversity coefficient (<i>KI</i>)	0.48645598194131	0.67184035476718
Syntactic complexity coefficient (<i>Ks</i>)	0.93271461716937	0.95709570957096
Speech coherence coefficient (<i>Kz</i>)	1.7471264367816	2.2307692307692
Index of exclusivity (<i>Iwt</i>)	0.76566125290023	0.83168316831683
Index of concentration (<i>Ikt</i>)	0.02784222737819	0.0099009900990099

All in all, the findings have proven the coefficient of speech coherence (*Kz*) is most distinctive; compare: ST *Kz* equals to 1.7471264367816 while TT *Kz* totals to 2.2307692307692 and exceeds the ST *Kz* for the amount of 0,4836427939874.

6. Translation studies textual analysis

Let us consider the dominant EDA for future Translation Studies analysis of the original and the translation from the viewpoint of pragmatic interference.

The loss of a loved relative is one of the main themes of Donna Tartt's novel *The Goldfinch*. There are numerous examples of this EDA in the first chapter of the novel. In one of them, the main character of the novel is having a dream in which he remembers his mother. The chain of EDA makes it possible to understand the psycho-emotional state of the main character of the novel. Excerpts below accurately trace the author's accentuation on the emotional and psychological state of the protagonist, who tries to convince himself that his mother survived the horrific terrorist attack. In these examples, the climax of EDA is manifested in two disclosures: explicitly and implicitly. Below is the screenshot to make it easier to visualize.

ST	TT
Thanks to my fever I had a lot of weird and extremely vivid dreams, sweats where I thrashed around hardly knowing if it was day or night, but on the last and worst of these nights I dreamed about my mother	Через лихоманку я пережив <u>безліч дивних і надзвичайно яскравих марень</u> , обливаючись потом, не знаючи, <u>котра година</u> , але в останню й <u>найгіршу з тих ночей мені наснилася мати</u>
at the sight of her I was paralyzed) [...] as much as I wanted to, I knew I couldn't turn around, that to look at her directly was to violate the laws of her world and mine	помітивши її, я заціпенів [...] хоч як мені хотілося обернутися, я знав, що подивитися на неї очі в очі було б порушенням законів її світу і мого світу
Sometimes, in the night, i woke up wailing	<u>Іноді я прокидався вночі, виючи</u>
I felt a petty, irrational pang of disappointment	я <u>почував легкий ірраціональний біль розчарування</u>
Boris never seemed annoyed or even very startled when I woke him, as if he came from a world where there was nothing so unusual in a nocturnal howl of pain	Борис ніколи не дратувався й навіть не дуже дивувався, коли я його будив так, ніби прийшов сюди зі світу, в якому прокидатися вночі, виючи від болю, не вважалося чимось незвичним
My mother's bedroom was dark and cool, and the faint, just-detectable smell of her perfume was almost more than I could bear	У спальні матері було темно й прохолодно, і ледь чутні пахощі її парфумів були чимось більшим, ніж я міг витримати
a million other heart-piercing sights	мільйона інших деталей, що розривали мені серце
The living room—normally so airy and open, buoyant with my mother's presence —had shrunk to a cold, pale discomfort, like a vacation house in winter	Вітальня — за нормальних обставин така свіжа й відкрита, життєрадісна завдяки присутності моєї матері — тепер здавалася холодною, блідою й незатишною, схожа на дачний будинок узимку
I could feel my heart beating	<u>Я чув, як гупає моє серце</u>
I missed her so much I wanted to die: a hard, physical longing, like a craving for air underwater	Я так сумував за нею, що мені хотілося померти: то була тяжка фізична туга, схожа на страждання без повітря під водою
Lying awake, I tried to recall all my best memories of her —to freeze her in my mind so I <u>wouldn't</u> forget her [...] drifting uneasily between dreaming and sleep.	Не в змозі заснути, я намагався пережити всі свої найкращі спогади про неї — заморозити її у своїй свідомості, щоб ніколи не забути про неї) [...] важко переходячи від дрімоти до сну

Figure 11: chosen EDA subtext. Citation from [18; 23].

Next, the key word generator can help to heighten the awareness of meaningful choices.

Block 1 is preconditioned by the norms of the English language of possessive pronouns placement and usage as well as the presence of definite article *the*, the allomorphic feature for Ukrainian.

Thus, Block 1 proves objective necessity.

In focus – Block 3. These are the lexemes that capture, so to say, EDA in a most illustrative way. This serves kind of a “double-check” for the translator at the post-editing and proof-reading stage.

We pay attention to main parts of speech (verb, noun, and pronoun, adjective) and/or their derivatives.

	ST	TT
B	her - 8; and - 7; the - 6; was - 6;	
BI		бути - 3; світ - 3;
Block 3	wanted - 2; had - 2; woke - 2; mother's - 2; like - 2; could - 2; where - 2; world - 2; around - 2; much - 2; night - 2;	виючи - 2; чимось - 2; серце - 2; матері - 2; вночі - 2; очі - 2;
Block 4		витримати - 1; міг - 1; ніж - 1; більший - 1; через - 1; парфумів - 1; чутні - 1; ледь - 1; прохолодно - 1; темно - 1; спальні - 1; незвичним - 1; вважати - 1; пахощі - 1; інший - 1; мільйона - 1; здавалася - 1; чув - 1; узимку - 1; будинок - 1; дачний - 1; схожий - 1;

Figure 12: key words generated from chosen EDA subtext. C

EDA key words of ST in the chosen subtext: *nouns* – mother, night, world; *verbs* – wanted, woke.

EDA key words of TT in the chosen subtext: *nouns* – матір, серце, очі, ніч (вночі), світ; *verbs* – виючи.

The discrepancies become obvious; they, however, kindle the avid interest. With a close look, the following translator's decisions can be noticed:

- For ST lexemes *wailing* and (nocturnal) *howl* the translator offered one correspondence *виючи*. Applying componential analysis, it becomes obvious that the English lexeme *to wail* has a sememe *weeping* with corresponding *ридання*, while the ST lexeme *howl* is rendered with its direct Ukrainian correspondence *вуми*. The ST fragment *Sometimes, in the night, i woke up wailing* might have easily be transformed with the relevant lexeme *Іноді я прокидався вночі, ридуючи*. This alternative choice would not have distorted EDA sequence or “communicative dynamics” (in M. Baker's terms), vice versa, it would have produced more diverse utterances in the exclusivity coefficient of the TT.

- The key word generator does not recognize word phrases such as *heart-piercing* as separate meaningful units. Therefore, it was omitted from the automatic calculation. The translator's choice is preconditioned by the allomorphic feature of English and Ukrainian, i.e. the absence of Participle constructions in the Ukrainian language. Thus, to sound fluent and natural, the ST Participial construction *a million other heart-piercing sights* could not have been translated other way than *мільйона інших деталей, що розривали мені серце*. Translator's choice is well justified.

- The ST phrase *to look at her directly* was reproduced by means of *подивитися на неї очі в очі* (back translation: *to look at her eyes to eyes*). This choice re-projected EDA verbalization means

more expressive in the translation than intended in the original of the chosen subtext. If frequent, such instances with additional translation markedness might serve a fruitful platform for researching translator's individual style and manner of translation.

- In the analyzed excerpt, *world* is a counterpart of EDA since it is metaphorized via psycho-emotional pangs of pain, e.g. *I knew I couldn't turn around, that to look at her directly was to violate the laws of her world and mine* (i.e. the worlds of dead and alive) and *world where there was nothing so unusual in a nocturnal howl of pain*. This is well-preserved, partially accentuated, *я знав, що подивитися на неї очі в очі було б порушенням законів її світу і мого світу and ніби прийшов сюди зі світу, в якому прокидатися вночі, виючи від болю, не вважалося чимось незвичним* correspondingly.

Despite the small number illustrated, these few samples show that NLP tools are profound resources for Translation Studies analysis. Along with linguistics-oriented approaches, they are useful at different stages of translation process:

(a) pre-translation ST analysis, original author's statistical profiling, defining the authors idiostyle via linguometry and key word generators, key lexical and stylistic features digital markup – along with translator's expertise and experience, – will, undoubtedly, navigate the professional translator and add up to the successful translation;

(b) translation process: NLP tools can help to identify dynamic and stable elements in a text; decide on strategies such as explicitation or compensation, what is more important, to be able to balance them; it makes lexical choice easier for the translator; mismatches would become a conscious translator's decision rather than mistake or negligence.

(c) post-translation: from the viewpoint of translation assessment and critical evaluation, NLP tools, as well as PTC, help to detect the holistic picture, to trace objective and subjective shifts of cohesion and coherence in TT. More emphasis might be placed on transitivity structure and ideational function of the text.

7. Acknowledgements

The project has been conducted within the scope of scholarly research topic “Application of modern technologies in information optimizing of the processes in natural language” at Lviv Polytechnic National University.

At the initial stage the project underwent the consultancy of Ihor Kulchytsky, to whom we express our gratitude.

The text was digitally marked-up by Petro Vus (graduate of Lviv Polytechnic National University).

The process was scrupulously supervised by the authors of the current article.

8. Conclusions

Results and findings support the idea that quantitative indicators have the potential to demarcate principal vectors of the analysis of EDA dominants and their correspondences in translation. This elaboration gives impetus to talk about the quantitative equality and response with minor fluctuation of corresponding units.

The study analysis was conducted on the material of a corpus of texts, digitally marked up, from the viewpoint of computational and corpus linguistics.

Applied was qualitative analysis, dominant in modern linguistic studies. This allowed us to trace a holistic statistics picture of the author and the translator.

The upcoming results help to draw the discrepancies of parallel profiling; some of the unavoidable (due to the difference in pair language systems), and avoidable (translator's choices).

The perspective is recognized if proceeding with the comparison of multiple translation reproductions of the same source text writing.

9. References

- [1] I. A. Bolshakov, A. Gelbukh, *Computational Linguistics: Models, Resources, Applications*, Instituto Politecnico Nacional, Dirección de Publicaciones, Tresguerra, 2004. URL: <https://www.gelbukh.com/clbook/Computational-Linguistics.htm>
- [2] *Corpus-based Language Studies: An Advanced Resource Book*, ed. T. McEnery, R. Xiao, Y. Tono, Routledge, 2006.
- [3] *Corpus-based Translation Studies: Theory, Findings, Applications*, ed. S. Laviosa, Rodopy, 2002
- [4] N. S. Dash and S. Arulmozi, *History, Features, and Typology of Language Corpora*, Springer: Springer Nature Singapore, 2018
- [5] N. Hrytsiv, I. Kulchytskyy, O. Rohach, *Quantitative Comparative Analysis in Parallel Translation Corpus: building author's and translator's statistical profiles: (a case study of Lucy Maud Montgomery)*, 2020 IEEE 15th International Conference on Computer Sciences and Information Technologies (CSIT), pp. 255–258 (2020) doi: 10.1109/CSIT49958.2020.9321893
- [6] S. Hunston, *Corpora in Applied Linguistics*. Cambridge: Cambridge University Press, 2002
- [7] D. Kenny, *Lexis and Creativity in Translation. A corpus-based study*. Routledge, 2001
- [8] E. Kim, R. Klinger, *A Survey on Sentiment and Emotion Analysis for Computational Literary Studies*. In: *Zeitschrift für digitale Geisteswissenschaften*. Wolfenbüttel 2019. doi: 10.17175/2019_008
- [9] V. Lytvyn, V. Vysotska, P. Pukach, Z. Nytrebych, I. Demkiv, R. Kovalchuk, N. Huzyk, *Development of the linguometric method for automatic identification of the author of text content based on statistical analysis of language diversity coefficients*, *Eastern-European Journal of Enterprise Technologies* ISSN 1729-3774 5/2 (95) 2018, pp. 16–28
- [10] V. Lytvyn, V. Vysotska, T. Hamon, N. Grabar, N. Sharonova, O. Cherednichenko, O. Kanishcheva (Eds.), *Computational Linguistics and Intelligent Systems. Proc. 4thInt. Conf. COLINS 2020. Volume I:Workshop*. Lviv, Ukraine, April 23-24, 2020, CEUR-WS.org, onlineO. Levchenko, O. Tyshchenko, M.Dilai. *Associative verbal network of the conceptual domain БИДА (MISERY) in Ukrainian*, *Proceedings of the 4th International Conference on Computational Linguistics and Intelligent Systems (COLINS 2020). Volume I: Main Conference, 2020, Pp. 106-120*
- [11] J. Munday, *A Computer-assisted approach to the Analysis of Translation Shifts*, *Meta*, XLIII, 4, 1998
- [12] S. Prabhume, A. W. Black, R. Salakhutdinov, *Exploring Controllable Text Generation Techniques*, *Proceedings of the 28th International Conference on Computational Linguistics, 2020*, pp.1–14. doi: 10.18653/v1/2020.coling-main.1, URL: <https://www.aclweb.org/anthology/2020.coling-main.1/>
- [13] N. Shakhovska, M. Medykovskyy (Eds), *Advances in Intelligent Systems and Computing III: Selected papers from the International Conference on Computer Science and Information Technologies, CSIT 2018, September 11-14 Lviv, Ukraine*. Springer: Springer Nature Switzerland (2019).
- [14] A. Soloshenko, *Emotions in legal fiction: conceptual metaphors and cross-domain mapping with ATLAS.ti*. Linguistics. Université de Strasbourg, 2016. URL: <https://tel.archives-ouvertes.fr/tel-01529768/document>
- [15] *Statistical NLP*, URL: <https://plato.stanford.edu/entries/computational-linguistics/#StaNLP>
- [16] J. Sinclair, *Corpus, concordance, collocation*, Oxford : Oxford University Press, 1991, 137 p.
- [17] D. Tartt, *The Goldfinch*, Little Brown Publishing, 2013, 771 p.
- [18] M. Tymoczko, *Computerized corpora and the future of translation studies*, *Meta*, 1998, 434, pp. 652–659.

- [19] M. Ulrych, The impact of multilingual parallel concordancing on translation. In *Practical Applications in Language Corpora*, ed. B.L. Tomaszczyk, and P.J. Melia, 1997, pp. 421–436. Lodz: Lodz University Press.
- [20] F. Zanettin, Parallel corpora in translation studies: Issues in corpus design and analysis. In *Intercultural Faultlines. Research Models in Translation Studies I: Textual and Cognitive Aspects*, ed. M. Olohan, 2000, pp. 105–118. Manchester: St. Jerome.
- [21] The Second Workshop on Human-Informed Translation and Interpreting Technology (HiT-IT 2019): Proceedings of the Workshop, Varna, Bulgaria, September 5-6, 2019. URL: <http://rgcl.wlv.ac.uk/hit-it2019/>
- [22] D. Tartt, Shchyhol. Perekład z anhliiskoi: Viktor Shovkun, Kharkiv: KSD, 2016, 816 stor.