ПӘНДЕР АРАЛЫҚ ҒЫЛЫМДАР
МЕЖДИСЦИПЛИНАРНЫЕ НАУКИ
INTERDISCIPLINARY SCIENCES

CENTRAL ASIAN SCIENTIFIC JOURNAL
ISSN 3135-3061
2026. No. 4 (32), Vol. 2. pp. 301–304
DOI: https://doi.org/10.5281/zenodo.24170221
Journal website: https://cajournal.kz

UDC 821.161.1:004.8

WHAT A WORD COUNT CANNOT READ: THE EMOTIONAL ARC OF CRIME AND PUNISHMENT MEASURED WITH THE LABMT DICTIONARY AND A READER'S INTERPRETATION

Sapayeva Danaiya
Zhaniya Aubakirova Author's School, Almaty, Kazakhstan

Sapayeva Danaiya — Grade 11 student, Zhaniya Aubakirova Author's School, Almaty, Kazakhstan.

© Sapayeva Danaiya, 2026

Abstract: Computer methods that measure the emotional arc of a novel by averaging the happiness of its words are now widely used. This study applies one such method, the labMT word list, to the English translation of Dostoevsky's Crime and Punishment by Constance Garnett, chapter by chapter, and compares the result with the author's own reading of the novel. The author reads the novel as the shape "fall", with guilt gradually consuming Raskolnikov, and the epilogue as light. The word count disagrees on both points at the level of chapters: the six parts of the novel stay within a narrow band (5.63-5.72), the epilogue is the darkest part (5.48), and no simple story shape fits well (the best correlation is 0.30). A closer reading changes the picture in part. Of three epilogue sentences the author chose as the lightest after seeing the chapter results, two score among the brightest in the novel (higher than 99% and 88% of its sentences), while the third scores among the darkest tenth; the mood of the last chapter rises from 5.11 in its first fifth to 5.80 in its last; and the words that carry her reading in the English translation, "renewed", "renewal" and "regeneration", are not in the dictionary at all. The low score of the epilogue comes mostly from negations and from words of its prison and illness setting, while its ending scores as light, as the reader saw.

Keywords: sentiment analysis, emotional arc, labMT, Dostoevsky, Crime and Punishment, digital humanities

1. Introduction

In 2016 Reagan and colleagues measured the emotional arcs of 1,327 stories from Project Gutenberg by sliding a window through each text and averaging the happiness of its words, and found that most arcs belong to six basic shapes: rise, fall, fall-rise ("man in a hole"), rise-fall ("Icarus"), rise-fall-rise ("Cinderella") and fall-rise-fall ("Oedipus") [1]. The happiness of a word comes from the labMT list, in which about 10,000 frequent English words were rated from 1 (saddest) to 9 (happiest) by human participants [2]. The method is attractive because it can be applied to any long text, but it reads words, not meaning.

The author found the Reagan study while looking for ways to measure literature with numbers, and chose Crime and Punishment, a novel she knows well, to see whether the method agrees with a reader. The question of this study is where a word count of the novel agrees with a careful reading and where it does not.

2. Hypothesis

After seeing the first chart of happiness by chapter, and before the chart was compared with the six shapes, the author wrote (18 October 2026) that the novel is closest to the shape "fall",

"because guilt and fear gradually consume Raskolnikov", but that the epilogue is light, "since love for Sonya becomes the beginning of his spiritual rebirth", even though the program scores it as dark (original in Russian). After seeing the comparison with the shapes, she chose (23 October) three epilogue sentences that she reads as the lightest, each with the Russian word that carries its meaning: "Их воскресила любовь" (воскресила, "resurrected"), "Но тут уже начинается новая история, история постепенного обновления человека" (обновления, "renewal") and "Они хотели говорить, но не могли" (хотели, "wanted"). In Garnett's translation these are "They were renewed by love", "But that is the beginning of a new story--the story of the gradual renewal of a man" and "They wanted to speak, but could not", so her key words appear there as "renewed", "renewal" (with "regeneration" in the same sentence) and "wanted".

3. Data and methods

The text is the English translation by Constance Garnett from Project Gutenberg (eBook #2554), downloaded on 29 October 2026 [3]. It was divided into its 41 chapters: six parts and two chapters of the epilogue. The happiness of each chapter is the average labMT score [2] of its words; as in the labMT method, words with neutral scores between 4 and 6 were left out. The chapter series was smoothed with a five-chapter moving average, standardised and compared, by Pearson correlation, with simple templates of the six shapes from [1] (a straight rise, a straight fall and half or one and a half sine waves). The three sentences chosen by the author were found in the translation and scored in the same way. To see how the mood changes inside the last chapter, the second chapter of the epilogue was divided into five parts of equal length. All calculations were done with a short Python script.

4. Results

4.1. The novel chapter by chapter

The six parts of the novel have almost the same average happiness, between 5.63 (Part II) and 5.72 (Part III) (Table 1). The darkest chapter is Part I, Chapter VII, the murder of the pawnbroker (5.35); the brightest are Part I, Chapters III and IV, built around the letter from Raskolnikov's mother (5.91 and 5.91), and Part VI, Chapter IV (5.89). The epilogue, at 5.48, is the darkest part of the book. None of the six shapes fits the smoothed series well: the closest are "Oedipus" ($r = 0.30$) and "fall" ($r = 0.21$); each opposite shape has the same r with the sign reversed, because its template is the negative of the other. The weak fit with "fall" depends on the epilogue: without the two epilogue chapters it becomes $r = -0.05$, and on the unsmoothed series $r = 0.09$. With 41 overlapping smoothed points these correlations are descriptive and were not tested for significance. By this measure the novel does not fall gradually; it stays nearly level and darkens at the end.

Table 1. Mean labMT happiness of the chapters in each part of Crime and Punishment (scale 1-9)

Part	I	II	III	IV	V	VI	Epilogue	Whole novel
Mean happiness	5.68	5.63	5.72	5.68	5.68	5.71	5.48	5.67

4.2. The epilogue sentence by sentence

To compare single sentences fairly, every sentence of the novel was scored in the same way (11,683 sentences with at least one scored word). Two of the three sentences chosen by the author are among the brightest: "They were renewed by love; the heart of each held infinite sources of life for the heart of the other" (7.32) and the closing sentence about "a new story" (6.62), higher than 98.8% and 87.8% of all sentences. The third, "They wanted to speak, but could not; tears stood in their eyes", scores 4.19, higher than only 10.3% of sentences, because the only scored words are "not", "tears" and "eyes" (Table 2). The words that carry her reading are invisible to the method:

"renewed", "renewal" and "regeneration" are not in the labMT list (nor are "resurrected" and "resurrection"), and "wanted" (5.08) falls in the neutral band that is left out.

Table 2. labMT scores of the three epilogue sentences the author reads as the lightest (novel average 5.67)

Sentence chosen by the author	Score	Words the method scored
"They were renewed by love..."	7.32	love 8.42, heart 7.22, infinite 6.44, life 7.32, heart 7.22
"But that is the beginning of a new story..."	6.62	beginning 6.04, new 6.82, story 6.48, story 6.48, story 6.48, world 6.52, new 6.82, life 7.32
"They wanted to speak, but could not..."	4.19	not 3.86, tears 2.46, eyes 6.24

Inside the last chapter the mood rises, with one dip: the five equal parts of Epilogue, Chapter II score 5.11, 5.33, 5.66, 5.54, 5.80. The first part, about prison life and Raskolnikov's illness, is among the darkest passages of the book; the fourth, which retells his feverish dream, falls back slightly; and the last part, which begins with the reconciliation with Sonya ("He wept and threw his arms round her knees"), is above the novel's average.

5. Discussion

A likely reason for the disagreement is that a chapter average treats every word in the chapter equally. The words that pull the epilogue furthest below the novel's average, counting how often each occurs and how far its score lies below the average, are "not" (69 times, 3.86), "no" (23 times, 3.48), "prison" (13 times, 2.06), "ill" (11 times, 2.42), "crime" (10 times, 2.20). The first two are negations, which the labMT list scores as unhappy whatever they negate (the third sentence the author chose is scored low partly for its "not"); the next are words of the prison and illness setting. The meaning the author sees in the epilogue is carried by a few words about change, and these are exactly the words the dictionary does not contain. When the text is read more closely, sentence by sentence and part by part, the method moves part of the way towards her reading: two of the sentences she chose are among the brightest in the novel, and the last chapter ends brighter than it begins, although the third sentence, about tears, is scored as dark.

Her first expectation, a gradual fall, is not supported: the six parts are almost level, and the darkest chapter, the murder, comes at the end of Part I rather than late in the novel. One reason may be that guilt and fear in Crime and Punishment are expressed through thoughts and situations rather than through consistently dark words, so the inner fall that a reader feels does not appear in the average of the words.

Limitations. The analysis uses one English translation, so the scores reflect Garnett's vocabulary as well as Dostoevsky's story. The labMT list was rated by modern American participants and contains about 10,000 words; many literary words are missing. Chapter averages differ in length, and the shape templates are simplified versions of those in [1]. Three sentences are an illustration, not a sample, and they were chosen after the chapter results were known.

6. Conclusion

Measured word by word, Crime and Punishment stays almost level and its epilogue is the darkest part. Measured more closely, the epilogue moves from darkness towards light, and two of the three sentences a reader found most hopeful are among the brightest in the book. Methods based on word lists capture the setting of a passage, and even its negations, better than its meaning, and in a novel about inner change the difference matters.

Data availability

The text and the labMT list are public [2, 3]. The chapter scores and the analysis script are available from the author or her mentoring team on request.

© Central Asian Scientific Journal, 2026

Author contributions and use of AI tools

The author chose this research direction from five proposed by her mentor, found the study of Reagan and colleagues used as the basis of the method, gave the reading of the novel quoted in Section 2, and selected the three epilogue sentences and their key words. The analysis was carried out and the text was drafted by her mentor with the help of a generative AI assistant (Claude, Anthropic). Every number was computed by the analysis script from the sources listed below. Before submission, the author checked the results herself: she recalculated by hand the score of the sentence "They were renewed by love..." from its five scored words in Table 2, (8.42 + 7.22 + 6.44 + 7.32 + 7.22) / 5, and obtained 7.32, matching the value in the paper.

Conflicts of interest

The author declares no conflicts of interest. No funding was received for this study.

References:

[1] Reagan A.J., Mitchell L., Kiley D., Danforth C.M., Dodds P.S. The emotional arcs of stories are dominated by six basic shapes // EPJ Data Science. 2016. Vol. 5. Art. 31. https://doi.org/10.1140/epjds/s13688-016-0093-1

[2] Dodds P.S., Harris K.D., Kloumann I.M., Bliss C.A., Danforth C.M. Temporal patterns of happiness and information in a global social network: hedonometrics and Twitter // PLoS ONE. 2011. Vol. 6, No. 12. e26752. https://doi.org/10.1371/journal.pone.0026752 (labMT 1.0 word list: https://hedonometer.org, accessed 29.10.2026).

[3] Dostoevsky F. Crime and Punishment / translated by C. Garnett. Project Gutenberg eBook #2554. https://www.gutenberg.org/ebooks/2554 (accessed 29.10.2026).