From the Tractatus Logico-Philosophicus to Later Wittgenstein: An NLP-Based Comparative Analysis

Andreiana Mihail[*]
Universitatea Bucuresti
andreiana.mihail@gmail.com

Silviu-Florin Gheorghe[*]
Universitatea Bucuresti
ghesil@gmail.com

Andrei Fotea
Universitatea Bucuresti
andreifotea@gmail.com

Liviu P Dinu
Universitatea Bucuresti
ldinu@fmi.unibuc.ro

Abstract

This study investigates the application of Natural Language Processing (NLP) methods to uncover linguistic and stylistic variations within the corpus of Ludwig Wittgenstein, a philosopher renowned for his complex and notional contributions. By analyzing works such as *Tractatus Logico-Philosophicus* alongside his later notes, manuscripts, and student-dictated lectures in Cambridge, we aim to identify significant distinctions in language use and conceptual framing. The corpus poses unique difficulties because of its diverse origins, including published works, personal notes, and collaboratively edited transcripts.

Using zero-shot NLP techniques, this exploratory/preliminary research aims to reveal patterns reflective of Wittgenstein's philosophical evolution and differences in text production methods. The results highlight the potential of computational approaches to enhance our understanding of complex, context-dependent philosophical writings, providing a possible path for further interdisciplinary investigations into linguistic and conceptual dynamics in the challenging body of work.

1 Introduction

Ludwig Wittgenstein was an Austrian philosopher (1889-1951) whose oeuvre focused mainly on logic, mathematics and language. His only book published during his lifetime was the 75 - page *Logisch-Philosophische Abhandlung (Tractatus Logico-Philosophicus)*, which, by Stanford Encyclopaedia of Philosophy "(. . .) addresses the central problems of philosophy which deal with the world, thought and language, and presents a 'solution' (as Wittgenstein terms it) of these problems that is grounded in logic and in the nature of representation." The *Tractatus* is written in a dry and aphoristic style, with no arguments, and consists of altogether 525 declarative statements, which are hierarchically numbered.

Posthumously, there were thousands of pages of manuscripts found - some of them not published even to this date - courses he held in Cambridge in the 1930s, notes and research reports written mostly in German, his native language. His major text is the *Philosophical Investigation* (Wittgenstein, 1968) which philosophy scholars consider the most important philosophy treatise of the 20th century. In its introduction, Wittgenstein admits the changes he made along the way, so that it can be understood "only by contrast with and against the background of my old way of thinking"(Wittgenstein, 1968), meaning in the *Tractatus Logico-Philosophicus* (Wittgenstein, 1921a). Many of the ideas developed in the *Tractatus* were criticized in the *Philosophical Investigations*, while others were later developed. The initial challenge of this project came from the *Yellow, Blue and Brown Books*, which were written after dictation by three of his students, Margaret Macdonald (Masterman), Francis Skinner and Alice Ambrose, during the lectures he held in Cambridge between 1932-1935. As one of the students, Alice Ambrose, admitted in the 1992 edition of Ludwig Wittgenstein's Lectures - Cambridge 1932 - 1935 *(The Yellow Book)* : "The difficulties I had finishing this edition of these notes taken between 1932 and 1935, a time when I attended Ludwig Wittgenstein's lectures, have multiplied because I didn't have access to other students' notes who took the same course, in order to check my notes, except for the year '34 -'35. (. . .) I rarely tried to make reference to some other texts already published. I also didn't attempt to follow the exact order of Wittgenstein's lectures." There were some fragments which were excluded, because of repetitions and digressions, because of the oral manner of the notes.

The *Yellow Book* together with the *Brown and*

---

[*] Equal contribution.

Blue Books marked the fundamental change in Wittgenstein's philosophical venture, moving from the exact mathematical and logical atomism towards a more approachable philosophical language. It is considered an important resource for understanding the development of Wittgenstein's ideas and the shift in his philosophical approach. After his death, there were many controversies about his intellectual heritage and the lack of control over the early and unchecked manuscripts he left. There are still debates over his intellectual testament and its handling by various people who embarked in this process.

Inspired by (Uban et al., 2021) we decided to look at the evolution of ideas in the writings of Ludwig Wittgenstein from the stylistic and semantic perspectives. The novelty of this research consists of performing a computer aided analysis of the work of Wittgenstein using existing proven tools, and trying to interpret the results without any active expectations. This could be relevant since he was one of the first scholars to discuss the statistical nature of language and the fact that the words are defined by their use.

As workflow, we wanted to see whether the texts would be grouped by the translator, by the language the text was written in, or by the year of creation, when analysed from a stylistic standpoint. We also wanted to know if comparing the texts using a pre-trained language model (PLM) leads to better results than the same analysis based on word distribution and bag-of-words (BOW) methods. The novelty does not consist of the tools or methods applied, but of using those tools on a specific corpus, with the clear goal of searching for new insights and information on the context of the writings.

2 Related Work

Stylometry & Authorship dentification Text analysis based on style and word content is very old, dating back at least to medieval times. For example, the *Donation of Constantine* is a forged Roman imperial decree in which, supposedly, emperor Constantine the Great transfers authority over Rome and the Western part of the Roman Empire to the Pope. The letter was contested, on stylistic grounds, by Lorenzo Valla (1597-1599). He realized that the language used in the text was inexistent in the 4th century, so that the document must be a forgery, a conclusion which was accepted. Valla's

argumentation was qualitative rather than quantitative. Quantitative methods started in the late 19th and early 20th century with the works of George Udny Yule (Yule, 1939, 1944) and George Kingsley Zipf (Kingsley Zipf, 1932). Nowadays, both stylometry and authorship identification are well developed fields. Stamatos (2009) advocates for the use of function words, because they are "used in a largely unconscious manner by the authors, and they are topic-independent". There are also other features of the language that can be used[1].

In our case, this class of methods may be hindered by the fact that the use of stop-words is overshadowed by the translators' approach, who added their own stylistic signature on top of the one of the original author. Rybicki (2012) shows that, in a translated text, stylometric methods tend to indicate the original author, rather than the translator. The author identification on texts written by multiple people can be an intriguing topic as seen in (Dinu et al., 2012a; Dauber et al., 2017).

Thematic Analysis When trying to analyse the concepts and the domain of a text, function words and translation become less important. Using uncommon words as the indicator of the meaning of a text and a measure of text similarity is an old idea and represents, for example, the basis of the TF-IDF method (Salton, 1986). Such methods have been long used in order to compare texts based on content but are known to fail when the text is abstract and lacking domain specific keywords. Another completely different way of performing thematic analysis on texts is provided by the pre-trained language models, also known as 'transformers' (Vaswani et al., 2017; Devlin et al., 2019). Because the sequence length processed by the current generation of transformers is limited, there are various methods to determine the semantic representation suggested for long texts.

We should also mention a notable study on the subject of Ludwig Wittgenstein and NLP, investigating the influence of Wittgenstein's philosophy on the evolution of mathematical approaches to language and the historical development of what is n:)) ow known as Natural Language Processing, *Wittgenstein's Philosophy of Language: The Philosophical Origins of Modern NLP Thinking* (Bain et al., 2022).

---

[1]for example the frequency of the '-ing' continuous form

3 Methodology

The texts within the corpus showed special characteristics: some were originally written in English, while others are translations from German into English. Additionally, certain texts were authored directly by Wittgenstein, and others are collections of notes compiled by various collaborators. Another distinction lies in the intended purpose of the texts—some were explicitly written for publication, whereas others were personal notes. The corpus has been thoroughly studied since its publication, but, of particular interest are the comparative studies that examine the writings as a whole, exploring intertextual relationships and thematic coherence across Wittgenstein's works.

Despite its importance, finding all the relevant texts proved to be a difficult task. We were only able to source most papers as image-based PDFs and we had to digitize and clean them up, both automatically - using python regexp-, and then manually for fine adjustments. The role of the regular expressions was to clean up the page numbers and other artifacts produced by the OCR. The quality of the resulting texts is not the highest, but it is adequate, considering the statistical nature of this study.

Some of the works contain a lot of drawings, and including them using a multimodal model (Baltrušaitis et al., 2018) would be an interesting direction for future research, but in this study, the drawings were simply left out.

The complete list of texts can be found in table 1.

A quantitative comparison of texts was conducted by analysing the distribution of function words, as their use is highly individualistic and can serve as the distinctive linguistic "fingerprint" of an author. According to previous studies rank distance is particularly well-suited for this type of analysis (Popescu and Dinu, 2008; Dinu et al., 2012b). Consequently, we computed the pair distances of the texts and also followed the methodology of the previous studies and used the words described in Mosteller & Wallace (Mosteller and Wallace, 1963). According to them, it is important to identify the function words, also known as "non-contextual" words in order to distinguish between authors, in our case to see if there are differences throughout Wittgenstein's works, especially the *Yellow, Brown and Blue Books* compared to the other texts. Unsurprisingly, using the NLTK stop-words

led to very similar results, considering that there is a considerable overlap between the two sets.

These methods provide valuable information on the linguistic style and authorial attribution, enabling a systematic exploration of textual patterns.

Beyond stylistic elements, the relations between the texts can also be examined based on their semantics. Wittgenstein often revisited particular subjects across different works or explored various themes in separate texts in no certain order. One objective of our study was, in fact, to determine to which extent the semantic relations can be identified using various methods and tools. We were particularly interested if the pretrained language models perform significantly better than the bag-of-words based methods, when applied on a corpus like this. We are aware of the diversity of the language models available but we decided to stick to BERT which is better understood and studied due to its age. In order to decide we performed two experiments:

We computed a classical cosine similarity based on the rare terms (in a TF-IDF sense) as the first measure of similarity between texts. Specifically, it is counting the number of occurrences of the uncommon words that appear in at least two works and compute the cosine distances (1 - cosine similarity) of those vectors. Then, based on those distances we performed word clustering.

Subsequently, we compared these results to the pre-trained language models (PLMs) based methods. PLMs have showed remarkably high performance in the zero-shot language understanding and classification tasks (Devlin, 2018; Gera et al., 2022). Another reason we chose these models is that they allow deeper semantic interpretations of texts without requiring extensive task-specific training data, making them suitable for analysing intricate philosophical writings.

Because the data of the corpus is limited, we had to use Large Language Models (LLMs), that already have knowledge of the language in general, in order to get a better semantic understanding. Also, the texts were too long to be embedded in a single chunk, so we had to split them. The acquired chunks are of consistent size and contain whole sentences. The maximum chunk length is 512 tokens. It resulted in a list of embeddings corresponding to each chunk. In order to represent the whole text we had to compute some kind of aggregate. We decided to use BERTTopic(Devlin et al., 2019) to create some clusters, in an unsupervised
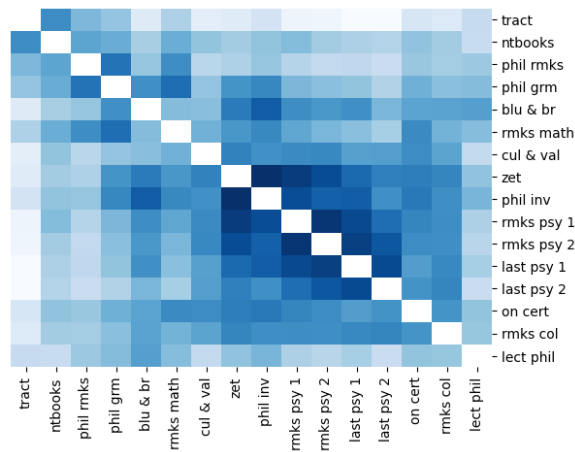
Figure 1: Similarity using the cosine similarity based on the number of function word occurrences. Darker means more similar.

Figure 2: Similarity dendrogram based on based on the number of function words occurrences.

Figure 3: Similarity based on TF-IDF.

way. The topics that cannot be clustered were discarded [2]. Consequently, we applied BERTopic to identify thematic clusters, each cluster containing a set of related text chunks. The resulting chunks were further grouped according to their original work and cluster. In this way, for each text, we obtained a vector of integers encoding the number of chunks of each topic that are present in each text. Then we computed the cosine similarity of those vectors as a measure of similarity between texts. Following normalization, we used the number of chunks associated with each topic as features to construct a semantic similarity matrix, allowing a quantitative assessment of thematic relations across different writings.

4 Experiments and Results

Function Word Similarity The pairwise rank distances based on the function words are plotted in Figure 2. The grouping and similarities are not a surprise, the texts being grouped by the period, mode of production and language.

An intriguing result was the lack of relatedness between *The Brown and Blue Book* and the *Philosophical Investigations,* even if the first was a preliminary study for the letter, raising a question to be addressed later on in a more exhaustive study we plan to put together with several philosophy scholars.

The other unexpected result was that the two volumes of *Last Writings on the Philosophy of Psychology* were not grouped together, how we thought

---

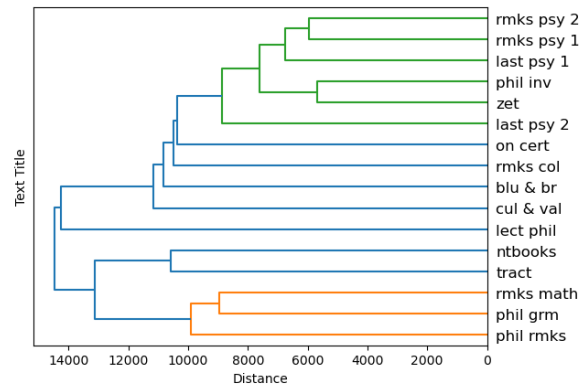[2]BERTTopic creates a special "-1" cluster representing the spurious data
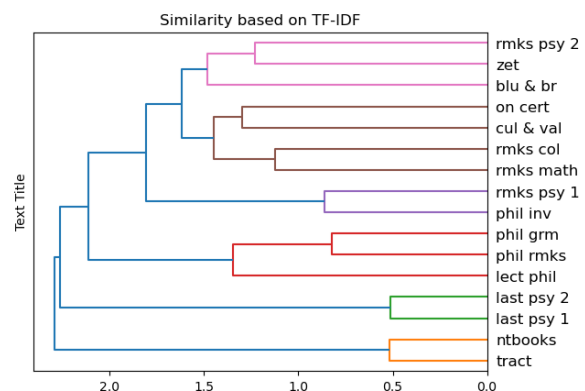
to be reasonable. This separation may be attributed to Wittgenstein's extensive archival material, which includes notes spanning different periods of his intellectual development. It is possible that the sources for these volumes emerged from distinct phases of his philosophical thought. As previously observed, Wittgenstein's work shows clear transitions, such as the shift from the *Notebooks* and the *Tractatus Logico-Philosophicus* to a later period characterized by the *Philosophical Investigations*. This later phase, often regarded as a new era in his philosophy, was preceded by the *Blue and Brown Books*, which served as preparatory texts for his evolving ideas. Otherwise the texts are reasonably and coherently placed on the distance map.

Long Word Similarity Calculating a similarity between texts based on the TF-IDF vectors (that are presumed to be the most specific to the topic, according to the efficiency of the language) would, ideally, cluster the texts based on their meaning. The texts are indeed reasonably and logically grouped, showing that this method can still be used for this type of task, even in the case of a small

Work	From	To	Short name	Lang
Tractatus (Wittgenstein, 1921b)	1921	1921	tract	G
Notebooks (Wittgenstein, 1916)	1914	1916	ntbooks	G
Philosophical remarks (Wittgenstein, 1930)	1930	1930	phil rmks	G
Philosophical grammar (Wittgenstein, 1933)	1933	1933	phil grm	G
The blue and the brown books (Wittgenstein, 1935a)	1933	1935	blu & br	E
Remarks on the foundations of mathematics (Wittgenstein, 1944)	1937	1944	rmks math	G
Culture and value (Wittgenstein, 1945)	1919	1945	cul & val	G
Zettel (Wittgenstein, 1948)	1929	1948	zet	G
Philosophical investigations (Wittgenstein, 1953)	1934	1935	phil inv	G
Remarks on the philosophy of psychology 1 (Wittgenstein, 1949b)	1946	1949	rmks psy 1	G
Remarks on the philosophy of psychology 2 (Wittgenstein, 1949c)	1946	1949	rmks psy 2	G
Last writings on the philosophy of psychology 1 (Wittgenstein, 1949a)	1948	1949	last psy 1	G
Last writings on the philosophy of psychology 2 (Wittgenstein, 1951a)	1949	1951	last psy 2	G
On Certainty (Wittgenstein, 1951b)	1949	1951	on cert	G
Remarks on Colour (Wittgenstein, 1950)	1950	1950	rmks col	G
Lectures on philosophy (Wittgenstein, 1935b)	1932	1935	lect phil	E

Table 1: List of works. Lang represents the original language.

corpus. Figure 3 shows the dendrogram that is resulting and the works are not grouped in any understandable way. One surprising feature is the distance between *Remarks on the Philosophy of Psychology 1 and 2* which can be explained by the numerous notes and writings of Wittgenstein produced in various periods of time.

Using Bert for Topic Detection   Not surprisingly, the results based on BERTTopic are a lot more relevant.

The results obtained using BERTTopic are presented in the figures 4 and 5.

In order to give the topics an intuitive meaning, we created a human-readable label for each topic. The labels were created as follows: the chunks associated to each topic have been first concatenated, forming pseudo-documents. Then, for each of those pseudo-documents we selected the best two unique terms, in the TF-IDF sense, to serve as cluster name. The source of some of the observed similarities is easy to spot in the correlation matrix. Also, it revealed the fact that some books are more focused on a topic (*Culture and Value)* while the topics in others are more spread out (*Lectures on Philosophy)*. Human intervention from scholars

with a good knowledge of Wittgenstein's work can improve this understanding, in an iterative semi-supervised manner.

Computing the pairwise distance between these representations led to Diagram 5.

5   Conclusions and Future Work

This study does not claim to be a philosophical analysis but a computational linguistic one. In conclusion, the Natural Language Processing (NLP) tools, both traditional and modern, offer valuable methods for analysing complex, philosophically dense corpora such as the one presented in this study, the oeuvre of Ludwig Wittgenstein. While producing certain conclusions remains challenging in the absence of a well-established baseline, these computational approaches serve as "intuition pumps" (Dennett, 1993), aiding in the discovery of novel insights and has good chances to become a useful tool for the philosophical approach and scrutiny of the texts. Furthermore, improved or manually chosen topic labels would significantly enhance interpretability and, eventually, the understanding of the fairly used concepts. Additionally, fine-tuning a large language model (LLM) specifically on philo-
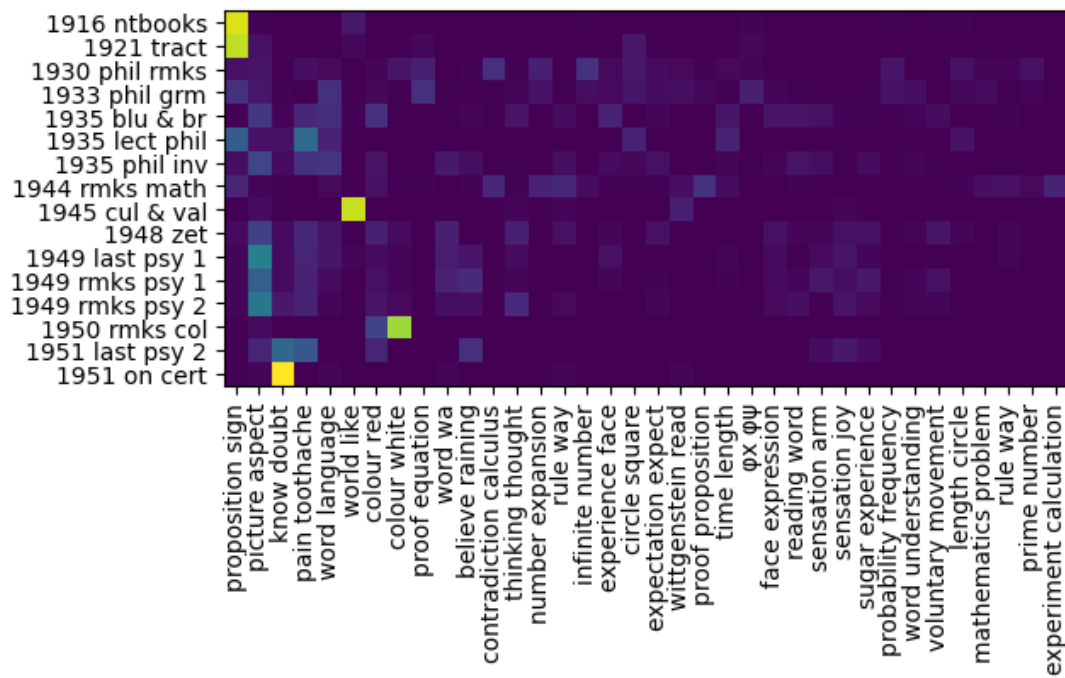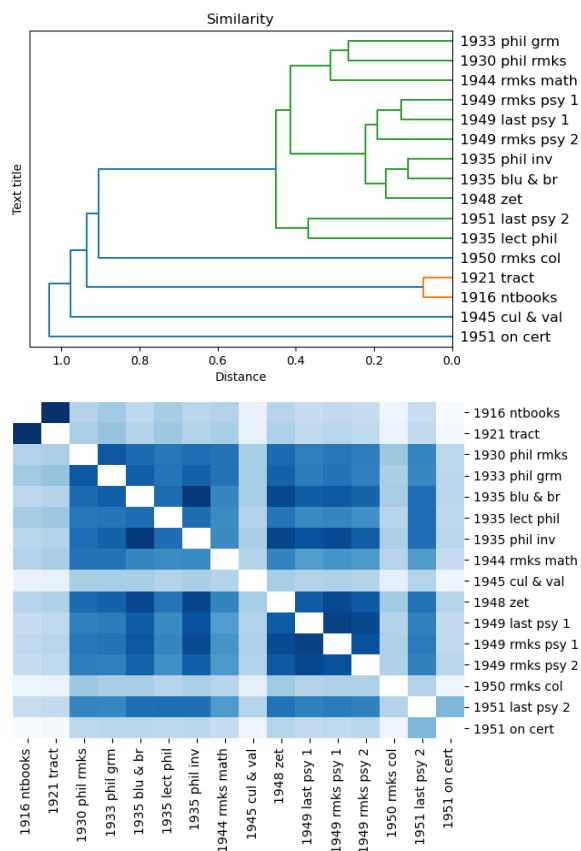
Figure 4: Topic by work

Figure 5: Topic similarity matrix and dendrogram based on the number of common topics, using LLM. Darker means more similar. Self similarity is not displayed.

sophical texts may improve the accuracy and relevance of topic detection. A particularly intriguing direction for future research is exploring the influence of language on text representations. Using multilingual models would be insightful in order to examine the proximity of the original German texts to their English translations in a language-independent semantic space. Although one might expect the translated texts to closely align with their German counterparts, the actual results are important and deserve empirical confirmation. Another approach would be comparative analyses between this corpus and the works of other contemporary philosophers and logicians, such as Kurt Gödel or Gottlob Frege, which can produce valuable insights into the capabilities and limitations of current semantic representation models. Conducting such a comparison of both original texts and their translations could provide a deeper understanding of the nuances of meaning that may be preserved, lost, or transformed through translation. A broader long-term research project could involve a large-scale study encompassing multiple philosophical works, including the exhaustive corpus from the Wittgenstein Archive at the University of Bergen (WAB), with the goal of mapping a history of ideas computationally and assessing how well such an approach aligns with the canonical history of philosophy. This would serve as an ambitious attempt

to quantify the evolution of philosophical thought using computational techniques.

6   Limitations

The lack of availability of a good quality text corpus presented a challenge, and some considerable effort went into producing and preprocessing the texts.

Automating the process of analysing and commenting on such texts is clearly beyond the scope of this study, as it presents significant challenges, as acquiring novel insights, particularly those not explicitly present in the training data, could fall within the domain of Artificial General Intelligence (AGI).

7   Ethical Statement

There are no ethical issues that could result from the publication of our work. Our experiments comply with all license agreements of the data sources used. We make the contents of our package available for research purposes upon request.

8   Acknowledgements

References

Robert Bain, Andrew Festa, Ga Young Lee, and Andy Zhang. 2022. Wittgenstein's philosophy of language the philosophical origins of modern nlp thinking.

Tadas Baltrušaitis, Chaitanya Ahuja, and Louis-Philippe Morency. 2018. Multimodal machine learning: A survey and taxonomy. *IEEE transactions on pattern analysis and machine intelligence*, 41(2):423–443.

Edwin Dauber, Rebekah Overdorf, and Rachel Greenstadt. 2017. Stylometric Authorship Attribution of Collaborative Documents. In *Cyber Security Cryptography and Machine Learning*. Springer International Publishing.

Daniel Dennett. 1993. *Consciousness explained*. Penguin uk.

Jacob Devlin. 2018. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*. Association for Computational Linguistics.

Anca Dinu, Liviu P Dinu, Alina Resceanu, and Ion Resceanu. 2012a. Some issues on the authorship identification in the apostles' epistles. In *Workshop Organizers*, page 18.

Liviu P Dinu, Vlad Niculae, and Octavia-Maria Şulea. 2012b. Pastiche detection based on stopword rankings. exposing impersonators of a romanian writer. In *Proceedings of the Workshop on Computational Approaches to Deception Detection*, pages 72–77.

Ariel Gera, Alon Halfon, Eyal Shnarch, Yotam Perlitz, Liat Ein-Dor, and Noam Slonim. 2022. Zero-shot text classification with self-training. In *EMNLP*.

George Kingsley Zipf. 1932. *Selected studies of the principle of relative frequency in language*. Harvard university press.

Frederick Mosteller and David L Wallace. 1963. Inference in an authorship problem: A comparative study of discrimination methods applied to the authorship of the disputed federalist papers. *Journal of the American Statistical Association*, 58(302):275–309.

Marius Popescu and Liviu P Dinu. 2008. Rank distance as a stylistic similarity. In *Coling 2008: Companion volume: Posters*, pages 91–94.

Jan Rybicki. 2012. The great mystery of the (almost) invisible translator: stylometry in translation. In *Quantitative Methods in Corpus-Based Translation Studies: A practical guide to descriptive translation research*, pages 231–248. John Benjamins Publishing Company.

Gerard Salton. 1986. Recent trends in automatic information retrieval. In *Proceedings of the 9th annual international ACM SIGIR conference on Research and development in information retrieval*, pages 1–10.

Efstathios Stamatatos. 2009. A survey of modern authorship attribution methods. *Journal of the American Society for information Science and Technology*, 60(3):538–556.

Ana Sabina Uban, Cornelia Caragea, and Liviu P Dinu. 2021. Studying the evolution of scientific topics and their relationships. In *Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021*, pages 1908–1922.

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Ł ukasz Kaiser, and Illia Polosukhin. 2017. Attention is All you Need. In *Advances in Neural Information Processing Systems*. Curran Associates, Inc.

Ludwig Wittgenstein. 1916. *Notebooks 1914-1916*. Blackwell, Oxford.

Ludwig Wittgenstein. 1921a. Tractatus logico-philosophicus.

Ludwig Wittgenstein. 1921b. *Tractatus Logico-Philosophicus*. Wilhelm Ostwald's Annalen der Naturphilosophie, Leipzig.

Ludwig Wittgenstein. 1930. *Philosophical Remarks*. Blackwell, Oxford.

Ludwig Wittgenstein. 1933. *Philosophical Grammar*. Blackwell, Oxford.

Ludwig Wittgenstein. 1935a. *The Blue and Brown Books*. Blackwell, Oxford.

Ludwig Wittgenstein. 1935b. *Lectures on Philosophy*. Blackwell, Oxford.

Ludwig Wittgenstein. 1944. *Remarks on the Foundations of Mathematics*. Blackwell, Oxford.

Ludwig Wittgenstein. 1945. *Culture and Value*. Blackwell, Oxford.

Ludwig Wittgenstein. 1948. *Zettel*. Blackwell, Oxford.

Ludwig Wittgenstein. 1949a. *Last Writings on the Philosophy of Psychology, Volume 1*. Blackwell, Oxford.

Ludwig Wittgenstein. 1949b. *Remarks on the Philosophy of Psychology, Volume 1*. Blackwell, Oxford.

Ludwig Wittgenstein. 1949c. *Remarks on the Philosophy of Psychology, Volume 2*. Blackwell, Oxford.

Ludwig Wittgenstein. 1950. *Remarks on Colour*. Blackwell, Oxford.

Ludwig Wittgenstein. 1951a. *Last Writings on the Philosophy of Psychology, Volume 2*. Blackwell, Oxford.

Ludwig Wittgenstein. 1951b. *On Certainty*. Blackwell, Oxford.

Ludwig Wittgenstein. 1953. *Philosophical Investigations*. Blackwell, Oxford.

Ludwig Wittgenstein. 1968. *Philosophical investigations. Wittgenstein*. Macmillan.

G Udny Yule. 1939. On sentence-length as a statistical characteristic of style in prose: With application to two cases of disputed authorship. *Biometrika*, 30(3/4):363–390.

George Udny Yule. 1944. *The Statistical Study of Literary Vocabulary*. CUP Archive.