Coreferential naming of characters in the book series "*The Witcher*"

Abstract

Co-reference resolution is an increasingly important area of research interest within NLP and training NLP-models on fiction texts could improve its accuracy. Previous studies have discussed the problem of co-reference resolution in journalistic texts in Russian (Dobrovolskii et al., 2022) and English (Chen et al., 2017) and have developed special NLP-models for that aim (Joshi et al., 2019). However, there is a lack of studies that cover the co-reference in a fiction, especially fantasy, texts using AI-trained models (Han et al., 2021). Co-reference resolution in fantasy texts has been largely neglected, so this study is intended to fill this lacuna. The purpose of the current study is to identify key features in characters’ co-reference naming in the dialogues and the author’s speech in “*The Witcher*” novel series. The current research employs quantitative and qualitative analysis, NLP methods and comparative analysis. The current paper might result in more accurate anaphora resolution in non-fiction texts using NLP-models and in better understanding of key features in co-reference in the dialogues and the author’s speech.

**Keywords**: Co-reference resolution, anaphora resolution, non-fiction texts, fantasy texts, “The Witcher”, natural language processing

Introduction

The issue of co-reference resolution is a fundamental challenge in natural language processing. In literary texts, which are the focus of this study, authors often employ literary devices, including those for co-referential references. Analyzing these devices can enhance the accuracy of co-reference resolution not only in other literary texts but also in conversational speech. Additionally, with the release of the Netflix series based on "The Witcher" saga and news about the upcoming game from CD Projekt Red set in the Witcher universe, the popularity of the original books is on the rise. This surge in interest underscores the significance of exploring co-reference patterns within "The Witcher" novels in the Russian language, offering valuable insights into how characters and entities are referred to and linked throughout the narrative.

The unique nature of literature poses challenges for co-reference resolution algorithms due to the presence of ambiguous pronouns, indirect references, and complex relationships between characters. By delving into the co-reference characteristics within "The Witcher" series in Russian, this study aims to shed light on the specific strategies employed by the author to establish and maintain referential continuity. Understanding these strategies can contribute to the development of more sophisticated co-reference resolution models capable of handling the intricacies of literary works.

Moreover, the current cultural resurgence surrounding "The Witcher" franchise, fueled by the multimedia adaptations and upcoming releases, underscores the relevance of studying co-reference within this context. By examining how co-reference operates within the rich narrative of "The Witcher" novels, this research seeks to provide insights that can benefit both the field of natural language processing and the broader appreciation and analysis of literary works in the Russian language.

In conclusion, this research endeavors to explore and analyze the character co-reference features present in "The Witcher" novel series in Russian, drawing attention to the nuances of referencing within complex literary narratives and their implications for computational linguistics and literary studies amid the renewed interest in the Witcher universe.

The research aimed at describing the peculiarities of using character co-referential expressions in the author's speech and dialogues covers a wide range of issues, including not only theoretical aspects of co-reference but also practical techniques for resolving co-reference using NLP models. The primary focus is on analyzing texts in the Russian language to identify the unique strategies employed by the author in establishing connections between characters. The research tasks involve studying existing materials on co-reference, analyzing methods of co-reference resolution using NLP models, collecting and annotating textual material, identifying errors in the model's operation, highlighting the speech styles of characters, and providing a detailed description of the peculiarities of using co-referential expressions in the author's speech and dialogues. This research approach will not only enhance the understanding of coreference processes in literary texts but also contribute significantly to the advancement of language analysis and processing methods employing modern technologies.

Literature Review

Proper names in literature have long been a subject of study for many linguists. A unique onomasticon can reveal certain characteristics of the group that uses it, for example, by correlating them with real-world groups of people. For some hobbits, Professor Tolkien wanted to use simple English names to emphasize the "Englishness" of this people, particularly in the case of the Shire. (Tolkien, 1944)

In addition to borrowing names from the real world, authors often create their own unique onomasticon. In literature, allusive names are often used to characterize their bearers in one way or another. For example, in "*The Lord of the Rings*", the name *Frodo* is derived from Old English *frōd*, meaning wise or prudent (Sweet, 1897), while the name Dumbledore in "Harry Potter," *Albus*, refers to the white color of his hair (from Latin *albus*, meaning *white*) (Gibka, 2019). In the Russian language, there are several ways of creating such anthroponyms: **suffixation** (often adding suffixes to the base to form surnames: *Трупов, Хлестаков, Собакевич*), compounding of bases (*Стародум, Щелколобов*), **substantivation** (*Солёный*), and **contamination** (*Пешеморепереходященская*). (Krylova, 2016). Such names called **allusive names**, they are based, in turn, on allusions that scholars studying this topic classify according to various criteria. For example, Solovyova's classification is based on the markedness of allusions in the text: allusive anthroponyms that **have markers of comparison** in the text and those **that do not** (Solovieva, 2004, p. 14). Additionally, a classification can be based on the characteristics used for comparison, as exemplified in Tsyrenova's work, where comparisons are made based on **external features** (hair, height, clothing, physique, etc.), **personal qualities** (habits, character), and **actions of the character** (behavior, deeds) (Tsyrenova, 2010, p. 14).

Another way of creating an anthroponym can be considered as a compound anthroponym, i.e. an anthroponym consisting of several types of onyms, for example, "The Witcher Saga" as a reference to the medieval tradition of naming where a personal name of a character is combined with a toponym of the place they came from (*Geralt of Rivia, Triss Merigold of Maribor*).(Tóth, 2022)

The typology of co-referential onyms can be classified into several categories based on their linguistic properties. These categories may include **pronominal onyms** such as personal pronouns (e.g., *he, she, it*), demonstrative pronouns (e.g., *this, that*), and possessive pronouns (e.g., *his, her*); **lexical onyms** such as proper nouns (e.g., *John, London*) and common nouns (e.g., *doctor, city*); and **nominal anaphors** (e.g., *the aforementioned, the former*). Additionally, these onyms can be examined in terms of their syntactic and semantic functions, as well as their discourse-level roles. Understanding the typology of co-referential onyms is crucial for comprehensive linguistic analysis and natural language understanding (Bach & Partee, 2004).

Attempts to describe anaphoric relationships have been made by linguists since the 1980s within the framework of generative grammar. These efforts led to the development of the **Binding Theory**, which has also evolved over the years. For instance, E. Reuland, in his theory of binding primitives, relies on functional explanations of coreference processes, one of which is the principle of economy (Reuland, 2001). When establishing anaphoric relationships, the speaker may refer to the syntactic, semantic, and discourse components of the text, and each of these levels can be positioned on a scale of economy from syntactic, the most "economical" in terms of resource expenditure, to discourse, the most "costly". According to Reuland's research, this suggests that speakers tend towards the least costly level of language when creating anaphors.

Other scholars, however, approached the issue of coreference from a discourse perspective. In this view, there were three main theories: **centering theory, accessibility theory, and the expectation hypothesis**.

**Centering theory** focuses on more local coreference, considering anaphora as a means of discourse coherence. Accordingly, discourse fragments with coreference are seen as more cohesive than those without it. Only one main linking element can exist between referents, and algorithms based on this theory describe only the anaphoric relationships of this element (Walker, 2007).

According to **accessibility theory**, the choice of anaphora is based on the accessibility of the referent in the speaker's memory, meaning that the more familiar or salient the object, the less detailed designation is required when mentioning it again. In contrast to centering theory, accessibility theory considers all possible anaphoric relationships between referents and places the speaker's attention at the center (Ariel, 2001).

The **expectation hypothesis** focuses on the addressee, assuming that the interlocutor immediately understands the referent being discussed, which accelerates the processing of anaphoric expressions upon their subsequent mention (Arnold, 2008).

These theories are of interest to us because most NLP models for co-reference resolution are based on this discursive approach. Over the years of co-reference resolution research, four different methods have been employed. Each subsequent method was built upon its predecessor, representing its advancement and elaboration. In chronological order, the following approaches to co-reference resolution can be distinguished: **Mention-Pair** models, **Mention-Ranking** models, **Entity-Based** models, and **Latent Structured** models.

**Mention-Pair** models represent the most basic and straightforward form employed in coreference resolution. They consider a pair of mentions at a time, along with the characteristics of each mention, and assign a binary outcome (Denis & Baldridge, 2007; Ng & Cardie, 2002; Soon et al., 2001).

**Mention-Ranking** models address the most apparent limitation of Mention-Pair models, which does not consider dependencies on other antecedent candidates, by simultaneously ranking and establishing a link only with the most highly ranked antecedent (Rahman & Ng, 2009; Yang et al., 2003).

However, Mention-Ranking models lacked the capability to determine when not to merge clusters. Errors arise in the resolution of merging mention clusters due to the significance of transitivity. **Entity-Based** models offer a knowledge classification approach to make informed decisions. This is also realized in Entity-Mention models and Cluster-Mention models, with the latter demonstrating significant improvements (Stoyanov & Eisner, 2012; Luo et al., 2004).

Following the previous models that attempted entity matching, **Latent-Structure** models emerged. These models differ from the preceding ones by focusing not on iteratively creating agglomerative clustering, meaning the grouping of similar elements, but on constructing a tree-like structure from which sections related to the same entity can be extracted (Marcoulides & Moustaki, 2014; Martins et al., 2019; Wu, 2022).

In recent years, many applications have transitioned from simple machine learning to deep learning. This shift was made possible by advancements in hardware that enable the creation of complex neural networks. Architectures and methods have also been developed to handle and utilize large volumes of data. A significant development in the field of deep learning was the introduction of word embedding concepts, followed by other language representation techniques (Mikolov et al., 2013).

**OntoNotes** is the standard for many datasets and widely used dataset for coreference resolution tasks. It consists of texts from various genres such as news articles, web texts, and religious texts, spanning seven genres and annotated in multiple languages including English, Spanish, and Chinese (Weischedel et al., 2012). It serves as a common benchmark for training and evaluating coreference resolution models. The dataset's features, like metadata on each text's genre and the authorship of each lemma, are integrated into many publicly available models. OntoNotes serves as a foundation for other datasets, like the FantasyCoref fantasy text corpus annotated following OntoNotes guidelines (Han et al., 2021). However co-reference resolution tasks are typically conducted on non-fiction datasets, “unlike non-fictional texts, referents in literary texts can be interpreted quite differently, depending on which point of view that the annotator takes (e.g., which character’s point of view, which part of the story that the reader is at)” (Han et al., 2021, p. 24).

Thus, the current work is aimed at filling the gap in co-reference resolution in the context of fantasy texts.

**Methods**

As the material for current paper I have chosen first 7 books of “The Witcher” novel series translated by E.P. Weisbrot. The selection of material is driven by the fact that these books are part of the main book series and are written by the creator of the series, Andrzej Sapkowski, and by the fact that one translator worked on these books while another translator worked on the subsequent ones, with differences in style and language between different translators potentially leading to unwanted errors in processing using a language model.

Initially, the compilation of a corpus from the specified books was necessary. Utilizing Python, I segmented each book into chapters, which were then extracted into 186 individual files (non-annotated text, NLP-metrics and annotated text), varying in content from 224 to 20,694 words and from 19 to 2,037 sentences. This process was implemented to mitigate the hardware load and time required for running the model while maintaining the quality of annotation results.

For the annotation phase, a Bert-based end-to-end model interface, developed by G. Gutnik for Dialogue-2021 (Gutnik, 2021). was employed. Bert stands as a standard in NLP-models and is commonly utilized for co-reference resolution. The specific model used for Russian annotation features a user-friendly interface for visualizing results and does not impose excessively high hardware requirements. Moreover, this model offers two distinct weights that will be compared within the scope of my study.

Upon the conclusion of the research, an analysis of the model's annotations will be conducted. Initially, a comparative evaluation of the two model weights will be undertaken to determine the most accurate option for fantasy texts. Furthermore, an investigation into the errors made by the model will be carried out to identify the underlying causes of such inaccuracies. Lastly, an examination of the methods through which co-reference is established will be performed, comparing the co-reference techniques utilized in characters' dialogues and the author's narrative.

**Results**

Within the raw annotated texts, I organized an excel table, where all the anaphors where divided by entities, they referred to. Also in the table, for each anaphora, its type and whether it was used by the author or the character is highlighted.

Preliminary findings indicate distinctions in co-reference usage between the author's narrative and the characters' speech. In characters' dialogues, the presence of co-references is contingent upon their knowledge and may be limited, whereas the "omniscient author" employs all possible co-references associated with a given entity (Han et al., 2021).

Furthermore, it seems that syntactic pronominal anaphors are the most dominant type of anaphora, with allusive anaphors emerging as a closely trailing counterpart in terms of prevalence.

It is anticipated that the model's precision will surpass that of non-fiction texts due to the richer context available for co-reference analysis in fantasy texts. However, the recall may diminish due to the abundant literary tropes within names, potentially leading to misinterpretations during analysis.

Additionally, it is expected that a significant number of errors will be linked to allusive names specifically, posing a challenge for accurate identification within the text.

**Conclusion**

In conclusion, it can be asserted that the aim of this study has been accomplished. The current investigation has elucidated specific regularities in the utilization of co-reference in fictional texts. Moreover, the analysis of the "The Witcher" book series material has substantiated the presence of the "omniscient author concept" (Han et al., 2021), wherein the author employs a comprehensive array of anaphors, while characters employ them based on contextual cues and prior knowledge.

Subsequently, it is conceivable to further this research by training a model on fantasy texts to ascertain the extent to which it may enhance analytical outcomes for both literary and non-fiction texts.

**References**

Ariel, M. (2001). Accessibility theory: An overview. In T. Sanders, J. Schilperoord, & W. Spooren (Eds.), *Human Cognitive Processing* (Vol. 8, p. 29). John Benjamins Publishing Company. https://doi.org/10.1075/hcp.8.04ari

Arnold, J. E. (2008). Reference production: Production-internal and addressee-oriented processes. *Language and Cognitive Processes*, *23*(4), 495–527. https://doi.org/10.1080/01690960801920099

Bach, E., & Partee, B. H. (2004). Anaphora and Semantic Structure. In B. H. Partee (Ed.), *Compositionality in Formal Semantics* (1st ed., pp. 122–152). Wiley. https://doi.org/10.1002/9780470751305.ch6

Chen, H. Y., Zhou, E., & Choi, J. D. (2017). Robust Coreference Resolution and Entity Linking on Dialogues: Character Identification on TV Show Transcripts. *Proceedings of the 21st Conference on Computational Natural Language Learning (CoNLL 2017)*, 216–225. https://doi.org/10.18653/v1/K17-1023

Denis, P., & Baldridge, J. (2007). Joint Determination of Anaphoricity and Coreference Resolution using Integer Programming. In C. Sidner, T. Schultz, M. Stone, & C. Zhai (Eds.), *Human Language Technologies 2007: The Conference of the North American Chapter of the Association for Computational Linguistics; Proceedings of the Main Conference* (pp. 236–243). Association for Computational Linguistics. https://aclanthology.org/N07-1030

Dobrovolskii, V., Michurina, M., & Ivoylova, A. (2022). RuCoCo: A new Russian corpus with coreference annotation. *COMPUTATIONAL LINGUISTICS AND INTELLECTUAL TECHNOLOGIES*, 141–149. https://doi.org/10.28995/2075-7182-2022-21-141-149

Gibka, M. (2019). Permanent Functions of Characters’ Proper Names in Harry Potter. *Journal of Literary Onomastics*, *7*(1), 44–52.

Gutnik, G. K. (2021). Opyt Adaptatsii Integral'nykh Modeley Razresheniya Koreferentsii Dlya Russkogo Yazyka [Experience of Adapting Integral Models for Coreference Resolution for the Russian Language]. *Dialogue-21*, Moscow. https://www.dialog-21.ru/media/5723/gutnikg142.pdf

Han, S., Seo, S., Kang, M., Kim, J., Choi, N., Song, M., & Choi, J. D. (2021). FantasyCoref: Coreference Resolution on Fantasy Literature Through Omniscient Writer’s Point of View. *Proceedings of the Fourth Workshop on Computational Models of Reference, Anaphora and Coreference*, 24–35. https://doi.org/10.18653/v1/2021.crac-1.3

Joshi, M., Levy, O., Zettlemoyer, L., & Weld, D. (2019). BERT for Coreference Resolution: Baselines and Analysis. *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, 5802–5807. https://doi.org/10.18653/v1/D19-1588

Krylova, A. G. (2016). Obshchie Priemy Sozdania Antronimov V Khudozhestvennom Tekste (Na Primere Proizvedenii A.P. Chekhova I Yu.V. Buidy). [General Principles of Creating Anthroponyms in Artistic Texts (Based on the Works of A.P. Chekhov and Yu.V. Buidy)]. *Rational and emotional aspects in the Russian language - 2016 Collection of Works of the International Scientific Conference*, 136–139.

Luo, X., Ittycheriah, A., Jing, H., Kambhatla, N., & Roukos, S. (2004). A Mention-Synchronous Coreference Resolution Algorithm Based On the Bell Tree. *Proceedings of the 42nd Annual Meeting of the Association for Computational Linguistics (ACL-04)*, 135–142. https://doi.org/10.3115/1218955.1218973

Marcoulides, G. A., & Moustaki, I. (2014). Latent Variable and Latent Structure Models. *Psychology Press.*

Martins, A. F. T., Mihaylova, T., Nangia, N., & Niculae, V. (2019). Latent Structure Models for Natural Language Processing. In P. Nakov & A. Palmer (Eds.), *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics: Tutorial Abstracts* (pp. 1–5). Association for Computational Linguistics. https://doi.org/10.18653/v1/P19-4001

Mikolov, T., Sutskever, I., Chen, K., Corrado, G., & Dean, J. (2013). Distributed Representations of Words and Phrases and their Compositionality (arXiv:1310.4546). arXiv. https://doi.org/10.48550/arXiv.1310.4546

Ng, V., & Cardie, C. (2002). Identifying Anaphoric and Non-Anaphoric Noun Phrases to Improve Coreference Resolution. *COLING 2002: The 19th International Conference on Computational Linguistics*. COLING 2002. https://aclanthology.org/C02-1139

Rahman, A., & Ng, V. (2009). Supervised Models for Coreference Resolution. In P. Koehn & R. Mihalcea (Eds.), *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing* (pp. 968–977). Association for Computational Linguistics. https://aclanthology.org/D09-1101

Reuland, E. (2001). Primitives of Binding. *Linguistic Inquiry*, *32*(3), 439–492. https://doi.org/10.1162/002438901750372522

Solovieva, M. A. (2004). Rol' Alluzivnogo Antroponima V Sozdanii Vertikal'nogo Konteksta (Na Materiale Romanov A. Merdok I Ikh Russkikh Perevodov) [The Role of Allusive Anthroponym in Creating Vertical Context (Based on the Novels of A. Murdoch and Their Russian Translations)] *[Author's abstract of dissertation for the degree of Candidate of Philological Sciences]*

Soon, W. M., Ng, H. T., & Lim, D. C. Y. (2001). A Machine Learning Approach to Coreference Resolution of Noun Phrases. *Computational Linguistics*, *27*(4), 521–544. https://doi.org/10.1162/089120101753342653

Stoyanov, V., & Eisner, J. (2012). Easy-first Coreference Resolution. In M. Kay & C. Boitet (Eds.), *Proceedings of COLING 2012* (pp. 2519–2534). The COLING 2012 Organizing Committee. https://aclanthology.org/C12-1154

Sweet, H. (1897). The Student’s Dictionary of Anglo-Saxon (*Oxford UP*).

Tolkien, J. R. R. (1944). Letter 76: To Christopher Tolkien, 1944 [Personal communication].

Tóth, V. (2022). Personal Names in a Medieval Context. *Helmut Buske Verlag*.

Tsyrenova, A. B. (2010). O Klassifikatsii Alluzivnykh Imen (Na Materiale Angliyskogo Yazyka) [On The Classification of Allusive Names (Based on the Material of the English Language)]. *Herald of Tomsk State Pedagogical University*, *7*, 13–19.

Walker, M. A. (Ed.). (2007). Centering theory in discourse (Reprint). *Clarendon Press.*

Weischedel, R., Pradhan, S., Ramshaw, L., Kaufman, J., Franchini, M., & El-Bachouti, M. (2012). OntoNotes Release 5.0 with OntoNotes DB Tool v0.999 beta*. BBN Technologies*. https://catalog.ldc.upenn.edu/docs/LDC2013T19/OntoNotes-Release-5.0.pdf

Wu, Z. (2022). Learning with Latent Structures in Natural Language Processing: A Survey (arXiv:2201.00490). arXiv. https://doi.org/10.48550/arXiv.2201.00490

Yang, X., Zhou, G., Su, J., & Tan, C. L. (2003). Coreference Resolution Using Competition Learning Approach. *Proceedings of the 41st Annual Meeting of the Association for Computational Linguistics*, 176–183. https://doi.org/10.3115/1075096.1075119