

The Literary Irony in the Works of Juliusz Słowacki

Anna Medrzecka

The Institute of Literary Research
The Polish Academy of Sciences
anna.medrzecka@gmail.com

Abstract

The purpose of this text to present the research project on the presence of various types of irony in the work of Juliusz Słowacki, one of the most important and influential poets of Polish Romanticism. This poet is known for his masterful use of various types of irony, which are unique in Polish literature. The nature and manner of its usage changes with the development of the poet's work. The first element of said project is to show the development of the use of irony and the artistic means that are used to achieve an ironic effect. I present different types of irony, my main interest being *literary irony* (Hammon, 1997).

When examining irony, I use not only recognized methods of traditional literary studies, but also methods introduced by digital humanities, with particular use of tools provided by the CLARIN consortium. In my paper I present the process of creating a digital corpus of texts by Juliusz Słowacki and its use in research. The first important step was to use the LEM tool to generate statistics of lemmas and tags within the entire corpus, individual texts and groups of texts. Next, I created a list of stylistic irony markers that can be compiled in such a way that they can be detected using CLARIN tools. In some cases, the tools need to be adjusted, but most often it is possible to set the currently available tools so that the expected results can help in search of the described irony indicators.

1 Introduction

The purpose of this text is to present research carried out as part of the doctoral project *Irony in the works of Juliusz Słowacki in the light of digital humanities research*. The research is aimed both at verifying whether the definitions and theories of irony established in the literature on the subject like e.g. (Szturc, 1992) (Hammon, 1997), can be supported by computer analysis, and at checking whether the computer analysis can provide a new point of view on the issue of irony and romantic irony. The project has a unique character due to its subject matter, which are poetic and literary texts. This makes it impossible to use tools typical for rhetorical irony analysis, such as the analysis of the tone of voice or the context of the statement. In addition, it raises the need to distinguish between elements of the ironic style – which is a subject of said research – and those elements that are characteristic of a literary text in general, which sometimes can be mistaken with the irony determinants.

The research is carried out with the use of presently available CLARIN-PL infrastructure, in the cooperation with Wrocław University of Science and Technology. Given its particular requirements, it constitutes a major contribution to the development of tools that can be used in other projects.

The research is carried out in several stages, described later in this article.

First, it was necessary to use philological analysis and textological knowledge to compile the corpus of Słowacki's works, including the creation of a list of available materials and the selection of appropriate versions. This stage has already been completed.

Then, selected texts were digitized and saved in files adapted for computer analysis.

This work is licenced under a Creative Commons Attribution 4.0 International Licence. Licence details: <http://creativecommons.org/licenses/by/4.0/>

The next step was to compile a list of determinants of the ironic style, especially those characteristic for Słowacki's work. The list has been preliminarily prepared, but is still being developed, because the conducted research often points out to unforeseen phenomena.

The current stage of the project is the selection, adaptation, testing, and development of digital tools that will best allow to explore the sought determinants of the ironic style in the research corpus. This part of work is still ongoing, so solutions proposed in this paper are likely to be developed still.

2 Juliusz Słowacki and his works – preliminary difficulties

The project is focused on the analysis of the body of work of one author, Juliusz Słowacki – one of the greatest Polish poets of the Romantic period. The preparation of the corpus and its adjustment to the analyses planned was the first important task that presented significant difficulties arising from the peculiar character of the author's artistic legacy.

Even as much as 75 percent of known and published works by Juliusz Słowacki are texts which remained unpublished until after the author's death. Afterwards, the poet's legacy ended in the hands of his mother, Salomea Bécu and his uncle, Teofil Januszewski. The latter handed over his late nephew's autographs to Antoni Małecki, a classical philologist, who in his few treatises on Polish literature mentioned Słowacki in a rather unfavourable way. Still, it is Małecki to whom we owe the first publication of many works left by the poet in autograph. Sadly, Januszewski did not prepare a catalogue of manuscripts delivered to the future publisher, so we are unable to determine the number of works Małecki decided to leave unedited for any reason. The unpublished works include not only letters or private notes of their author, including journal entries or a 'glossary' of Arab terms written down during his travels to the East, but also the overwhelming majority of poems, dramas, and narrative poems. Some of them present few textual problems – the versions conveyed in the manuscripts are uniform and coherent. However, a large portion of the legacy is made up of fragments of works or the so-called 'portions' later merged by the editors into the presently well-known works like the later parts of *Król-Duch* 'King-Spirit' or the so-called later songs of *Beniowski*. In case of the later works, separating individual portions would require massive editorial effort. It was therefore necessary to select the basis for texts put into the body of corpus. In most cases, I decided to base on the imperfect but still canonical Juliusz Kleiner's edition, both in terms of designating the beginning and end of individual texts, and the lection of the autographs. In case of texts having a newer critical edition, like e.g. Samuel Zborowski, published from manuscript by Marek Troszyński (Troszynski, 2017) numerous corrections were taken into account.

2.1 Corpus characteristic

In terms of corpus research, the varied status of texts (and the resulting different status of a word depending on its location) poses a significant problem. The works can be divided into:

1. **Texts published by the poet in his lifetime.** This group caused the fewest problems, because it can be assumed that the first printing reflects the author's intention, the status of a word is strong and unquestionable. Apart from obvious editing mistakes, some doubt can be posed in places by spelling and punctuation, as the comparison with manuscripts often shows that they may be a result of editorial intervention. For the purpose of the analyses described, spelling was standardized in all researched works in order to limit the number of incorrectly recognizes forms ('ign.') to the minimum and to obtain the possibly most reliable statistics. In researching irony spelling and punctuation are not without meaning. However, given the diversity of the material and the fact that it was impossible to obtain similarly reliable records in case of all texts, it was decided not to include this element in the analysis.
2. **Texts published after the poet's death, manuscripts missing.** In case of these texts there is a significant problem with assessing the reliability of the content (this group also includes the vast majority of the letters). Unfortunately, because it was impossible to verify the edited version against the manuscript original, it was decided for the sake of this analysis to accept critical editions.

3. **Texts published after the poet's death, surviving autographs.** Most often, these texts come in numerous versions, the relations of which is impossible to be determined today. In some cases, each version differs much from the others, in other the differences are minor. In classic editions these texts are often published with 'text variants'. In order to avoid falsifying the analysis results, it was decided to include text variants in the corpora, but without repeating fragments identical in several versions.

2.2 Text types and genres

Having determined the composition of the corpus formation, it became necessary to solve the problem arising from the diversity of genres of the researched texts. The corpus includes:

- 24 dramatic works and fragments
- 9 philosophical works
- 261 letters
- 30 narrative poems
- 25 prosaic works and fragments
- 209 poems

Here, it's worth adding that the texts differ much in terms of size. This results in a significant statistical imbalance. In order to avoid an extreme unbalance of texts in the corpus, the texts were divided into smaller files. This did not eliminate the disproportion completely, but still allowed to limit the falsification of data resulting from the use of the instances per million (i.p.m.) measurement (e.g. by the Kontext tool). I.p.m. shows the number of occurrences of a given word per million, related to the whole corpus or subcorpus (depending on the kind of analysis; corpus can consist of only one text). This measurement can be used to compare frequencies between corpora of different sizes. However, it should be remembered that although theoretically the relative frequency of the occurrence of a given word determined with the use of this tool will have a similar value for texts differing in size, in practice the lexical resource of a longer text is by definition much richer, so the results should be interpreted with great caution..

The base version of the corpus, containing all available texts in all existing versions, has been uploaded to the Kontext tool (Machálek, 2020) tool) which allows searching and creating frequency lists of words and tags. Kontext statistics show, that the tekst includes 1 180 127 word forms. The corpus is divided into subcorpora according to the generic characteristics described above.

2.3 Establishing corpus and subcorpora for analysis

In face of the apparent lack of balance in the corpus, a researcher has to solve numerous problems before attempting a proper analysis:

1. **How to select texts for analysis?** It was decided that wherever possible, all available texts should be included. It means including literary works as well as personal notes and letters.
2. **Is it methodologically justified to compare texts of different size and genre?** The problem of size was solved by dividing longer texts into shorter parts, and then using the i.p.m. measurement. It was also necessary because some of the used tools (e.g. WebSty) can only process files no larger than 20 Kb. The problem of genre was solved by dividing the corpus into subcorpora of texts with similar genre features. This allowed to single out features of irony characteristic to a given genre, not appearing in other subcorpora. For example, the use of verbs in 1st person singular in narrative poems may suggest its autothematic character, whereas in drama it is a genre feature.
3. **How to include changes related to the chronology of works in the analysis?** Texts from different periods differ in terms of subject and have certain specific stylistic features. The word *Duch* 'Spirit'

may serve as an example here. It is quite rare in works prior to 1843, whereas in the last years of the poet's life it becomes one of the most important lexemes, so much that it occupies the first place on the frequency list of the entire corpus. At the same time, given the nature of the legacy, it is not possible to precisely determine the order in which some of the later works were written. That is why while determining the analysis parameters it was necessary to balance them in a way that would make them possibly insensitive to elements related to the subject of a given work.

4. **How to treat erasures?** A word left in the manuscript is susceptible to removal, crossing out, being replaced with a new version. From today's point of view, it is impossible to describe relations of texts and versions. A text left in manuscript is truly an open work. So would it be responsible to treat manuscript words and published words as equal? Unfortunately, the only available solution would be to completely abandon analysing unpublished texts, and that would drastically reduce the research material. For this reason, it was decided to accept this imperfect decision and treat manuscripts and published works as equal sources, with relevant annotations.

3 In search for irony – current stage of research

The main subject of the research is the search for irony in the works of Juliusz Słowacki, with particular focus on romantic irony understood according to the theory by Friedrich Schlegel (Schlegel, 2019) and the classic publications for Polish irony studies like works of Włodzimierz Szturc and Zofia Mitosek. While elaborating on the concept of irony the researcher referred to literary studies, like the French *Ironie littéraire* (Hammon, 1997). The current state of machine text processing research on irony turned out relatively uninteresting, as it mostly focuses on small utility texts, like Twitter posts, or product reviews on online stores.

For this research it is important to determine the degree to which literary irony, discernible by a human, is visible in results of computer analysis. Definitely, this is not about creating a simple tool for irony recognition – it would seem that such a tool would not be possible at this stage, nor needed. In this project, computer analyses serve mainly as support for classical literary analyses. It is also a purpose of this research to find features of literary irony only visible via machine text processing. Already some interesting phenomena have been discovered as part of the work carried out so far, e.g. different types of temporal phrases used in texts recognised as ironical. The catalogue of features considered characteristic to an ironic text is currently being prepared for this purpose based on subject literature. The catalogue is still open. What's more, analyses conducted by a research team allow us to detect unexpected occurrences, and that is the greatest advantage of using digital tools.

Each occurrence is then described in a way allowing for its verification with the use of digital tools, thus making it possible for the researchers to verify characteristics of ironic style in the first place, and second to detect other stylistic elements pointing to an ironic reading that would be impossible to discern in traditional reading.

At this stage the catalogue of features is based mostly on the text material of the narrative poems. Detected ironic style features include:

1. **Repetitions**, broadly understood. Firstly, these are of course repetitions of one lexeme. Secondly, some fragments contain groups of different words related on a formative level, such as:
 - (1) Znałem... lecz, szczęściem ('fortune'), uleczona z żalu
Saffone, bardzo podobna do greckiej.
Ta sie, nieszczęściem ('misfortune'), kochała w Moskalu
A Moskał zginał na wojnie tureckiej
Ta poszła zabrać na warneńskim polu
Zwłoki, a uszy - w Konstantynopolu.
Smutna! Ubrana w kwiaty sympatyczne
Poszła nieszczesna ('unfortunate') na brzegi Marmora

2. **Proper nouns and foreign words on rhyme positions.** The rhyme position is naturally salient; emphasizing it by the use of a foreign word, one that naturally draws attention, provides yet another element of parabasis.
3. **Stylistic excess**, namely numerous appearance of words similar in meaning, emotional character, or register, in a limited fragment of a text. Unlike the simple repetitions, this may include various word forms and lexemes, and of very different meaning, only building up the ambience to the point when it becomes impossible to take seriously.

Bo i tu — i tam — za morzem — i wszędzie,	'Because here - and there - and everywhere
Gdzie tylko poszle przed sobą myśl biedna,	Wherever I put my poor thoughts
Zawsze mi smutno, i wszędzie mi jedno;	I'm always sad and always I feel the same
I wszędzie mi źle — i wiem, że źle będzie.	And I always feel bad - and I know I will feel bad.
Wiec już nie myślę teraz tylko o tém,	So I have no thoughts but this one
Gdzie wybrać miejsce na smutek łaskawe,	Where I can find a good place for my sadness
Miejsce, gdzie żaden duch nie traci lotem	A place no spirit would touch
O moje serce rozdarte i krwawe;	My broken and bloody heart
Miejsce, gdzie księżyc przyjdzie aż pod ławę	A place where a moon will come
Idać po fali...	On the waves...'

4. **Juxtaposed fragments differing in register**, such as comical fragments and fragments full of pathos. This feature of irony is explicitly called by the poet in several places, it is also characteristic for romantic irony, in which the elimination of language transparency is of particular importance.
5. **Autothematism, “writing about writing”, parabasis** – classic elements related to irony. This one is a higher-rank phenomena than those described above; here expressed by the accumulation of words related to writing ('I write', 'I describe', 'I say', 'I listen', 'I see', 'I sing', 'poet', 'poetry', 'song', etc.). In case of verbs, the ones appearing in 1st person singular (especially in narrative poems) are particularly important, their use in an author-narrated text has no other meaning than parabasis.

Naturally, the phenomena described above, are not invariably signs of irony; the purpose of this research is to verify the degree to which their appearance and co-appearance may point to an ironic reading. The examples above do not form a complete list of phenomena suggesting the use of irony in a literary text.

4 CLARIN tools

In order to verify the aforementioned features in texts from the corpus tools provided by the CLARIN consortium are being used.

4.1 Inforex

The tool to create annotated text corpora was used to gather all works and provide basic annotations with information about e.g. possible text variants (Marcińczuk, Oleksy, 2019).

4.2 LEM (Literary Exploration Machine)

LEM (Maryl, Piasecki, Walkowiak, 2017) is a versatile tool for text exploration, which in one place combines functionalities of other tools adding to their options. The main features of LEM include: lemmatization, part-of-speech tagging, generating custom wordlists and lemmatized texts. In the initial phase the tool was used to create the frequency list for the entire corpus, individual subcorpora, and singular texts. This allowed the researcher to perform preliminary stylometric analyses involving the frequency of words in individual subcorpora. The new “own categories” function allowed the text to be searched with reference to the above mentioned stylistic elements.

4.3 WebSty

The WebSty (Piasecki, Walkowiak, Eder, 2018) tool is used for stylometric analysis. As part of the project, the researcher is planning to use tools both for standard stylometric analysis, and with parameters precisely selected in terms of features described in the first part of this text. The planned analyses include the comparison of texts based on word lists pointing to autotelism – the list includes verbs like 'to write', 'to sing', 'to create', and nouns like 'poet', 'poetry', 'muse', 'song', etc.

4.4 NER

The NER tool turned out particularly useful while searching for autotelic fragments, as it allows users to analyse the occurrence of proper nouns and temporal words in a given text. The Liner2 (Marciniczuk, Kocon, Oleksy, 2017) tool made it possible to detect an additional feature of ironic texts, namely the occurrence of verbs differing in character on a limited space.

4.5 Słowosieć

plWordNet (Słowosieć) proved a particularly important tool in this research. The tool shows words grouped into sets of cognitive synonyms (synsets). It facilitates all analyses of tone, register, and search for synonyms and repetitions. It was also tested in terms of untypical word combinations or diminutive forms.

5 “Okno” ('Window') – an idea for the future

The characteristic feature of many stylistic elements related to irony is not just their occurrence (most of the features may carry other tropes or phenomena), but their frequency, or combination of several stylistic measures in one place. Hence the need for a tool that would not only check the occurrence of specified features in a whole text, but also detect their occurrence in one specific fragment. Such a tool is still in the concept stage, but its function could be based on numerous identical analyses for subsequent fragments, and later combining them in one statistic.

6 Current project state and perspectives for further research

Up to this point, all project stages focused on corpus forming tasks and allowed for a preliminary preparation of a catalogue of phenomena whose occurrence in text will be analyzed. The catalogue remains open also because new features appear as a result of examination research with the use of available tools. For it to be finished, it will be necessary to prepare additional functionalities for some of the elements of the CLARIN infrastructure; the current cooperation with the team at the Wrocław University of Science and Technology allows us to continuously introduce necessary improvements and functionalities.

The project is still underway so no final conclusions can be made. It seems, however, that the results obtained so far are promising and justify an optimistic outlook on future research in the designated direction, and the scientific methods prepared may be used in research on other corpora.

References

- Piasecki, M., Walkowiak, T., Eder, M. (2018, May). Open Stylometric System WebSty: Towards Multilingual and Multipurpose Workbench. In: *Selected papers from the CLARIN Annual Conference 2017, Budapest, 18–20 September 2017* (No. 147, pp. 145-158). Linköping University Electronic Press.
- Maryl, M., Piasecki, M., Walkowiak, T. (2018, May). *Literary Exploration Machine A Web-Based Application for Textual Scholars*. In: *Selected papers from the CLARIN Annual Conference 2017, Budapest, 18–20 September 2017* (No. 147, pp. 128-144). Linköping University Electronic Press.
- Marciniczuk, M. Oleksy, M. (2019). Inforex — a Collaborative System for Text Corpora Annotation and Analysis Goes Open. In *Proceedings of the International Conference on Recent Advances in Natural Language Processing, RANLP 2019, pages 711–719. Varna, Bulgaria. INCOMA Ltd.*
- Szturc, W. 1992. *Ironia romantyczna*, Wydawnictwo Naukowe PWN, Warszawa 1992.

- Troszyński, M. 2017. *Alchemia Rekopisu. Samuel Zborowski Juliusza Słowackiego* Wydawnictwo IBL, Warszawa 2017.
- Machálek, T. (2020). *KonText: Advanced and Flexible Corpus Query Interface*. In: Proceedings of LREC 2020, s. 7005–7010.
- Marcińczuk, M., Kocoń, J., Oleksy, M. (2017, April). Liner2—a generic framework for named entity recognition. In: Proceedings of the 6th Workshop on Balto-Slavic Natural Language Processing (pp. 86-91)
- Schlegel, F. 2019. *Fragmenty*, trans. Bartl, C. Wydawnictwo Uniwersytetu Jagiellońskiego.
- Hammon, Ph. 1997. *L'Ironie littéraire. Essai sur les formes de l'écriture oblique* Hachette, Paris.