Academic writing in the Baltic States: Introducing the Bwrite project

Anni Jürine

Djuddah A.J. Leijen
https://orcid.org/0000-0002-1630-7467
djuddah.leijen@ut.ee

Helen Hint
https://orcid.org/0000-0002-5588-112X
helen.hint@ut.ee

Jolanta Šinkūnienė
https://orcid.org/0000-0002-6803-8617
jolanta.sinkuniene@flf.vu.lt

Diāna Laiveniece
https://orcid.org/0000-0002-3107-2109
diana.laiveniece@liepu.lv

Christer Johansson
https://orcid.org/0000-0003-3715-3425
christer.johansson@uib.no

Nicholas Groom
https://orcid.org/0000-0001-8482-707X
n.w.groom@bham.ac.uk

In the project **Bwrite** (Academic Writing in the Baltic States: Rhetorical Structures through Cultures and Languages), we aim to address the lack of an empirically grounded holistic understanding of non-Anglophone writing traditions by mapping the academic writing traditions in the national languages of the Baltic States: Estonian, Latvian, and Lithuanian. We aim to achieve this by using machine learning and other computational methods (both quantitative and qualitative) for capturing writing tradition features at scale. By identifying and studying those features, we will not only create a body of knowledge on writing tradition(s) of the Baltic States, but the project will also provide a methodological basis for studying writing traditions elsewhere.

Keywords: Contrastive Rhetoric, Discourse Analysis, Corpus Analysis, Writing Traditions, Baltic States

Foreword

We would like to dedicate this article to the memory of Anni Jürine, our friend and colleague, who passed away on the 20th of April, 2021. Anni's dedication to the Bwrite project, her passion for the research and teaching of writing will continue to inspire us and everyone who has been in her presence.

1. Introduction

Ever since Kaplan's (1966) influential paper on cultural thought patterns in intercultural education, there has been an interest in how cultural backgrounds influence the way we communicate our ideas in writing. Often seen as the seminal paper in Contrastive Rhetoric, Kaplan's observations in intercultural classrooms gave rise to a multitude of studies contrasting aspects of different writing cultures primarily focusing on the challenges students encounter when writing in English as a second language motivated by EAP (English for Academic Purposes) research agendas. There has, however, been a relative absence of research focusing solely on non-Anglo-American oriented and non-Anglophone writing traditions. Some studies have attempted to characterise the major writing traditions of Europe in broad terms. For example, it has been suggested that the German writing tradition has a higher tolerance for digression (Siepmann, 2006) and the French writing tradition values style above all other considerations (Galtung, 1981). However, these claims about European writing traditions tend to be based on little or no empirical analysis and, as such, are not much more nuanced than the now notorious "doodles" that Kaplan (1996) used to characterize thought patterns across cultures (Severino, 1993).

The studies that have investigated European writing traditions empirically have done so from many different angles. Some empirical studies investigate writing at a micro-level focusing on specific linguistic features in texts, such as modal verbs, personal pronouns, boosters, and metaphors (Charteris-Black & Ennis, 2001; Mur-Dueñas & Šinkūnienė, 2016; Orta, 2010; Šinkūnienė & Olmen, 2012). Other empirical studies focus on the meso-level features of writing traditions, such as the application of self-promotion in Spanish journal articles (Martín & León Pérez, 2014) or on specific text types, such as conference abstracts in Ukrainian and Russian (Yakhontova, 2002). While these

studies certainly offer valuable insights into European writing traditions, the overall picture is scattered and sometimes even conflicting. This lack of coherence can be attributed to the fact that these studies focus on different levels of textual analysis or on different kinds of text, and generally use small amounts of data collected from specific populations operating in often very particular contexts. Often, the research on local writing traditions is published in other languages than English and, therefore, not available for wider audiences. As a result, we currently lack not only an empirically grounded holistic understanding of non-Anglophone writing traditions, but also a suite of modern methods (i.e., computational, quantitative, and qualitative) for capturing them at scale.

2. Bwrite project

In *Bwrite* (Academic Writing in the Baltic States: Rhetorical Structures through Cultures and Languages), we aim to address the above mentioned gap by mapping the academic writing traditions in the national languages of the Baltic States: Estonian, Latvian, and Lithuanian. Before we can embark on this endeavour, however, we will first need to identify the features that can be used to capture any given writing tradition. Once this model has been established, we will proceed to investigate how these features are manifested on the macro-level (the whole text), meso-level (paragraph level), and micro-level (sentence level) across a range of academic genres. In addition to building a body of knowledge on writing tradition(s) of the Baltic states, therefore, the project will also provide a methodological advancement in the form of a comprehensive analytical framework that can be used for investigating (lesser known) academic writing traditions elsewhere.

To determine the features that capture the Estonian, Latvian, and Lithuanian writing tradition(s), we start from studies conducted in the field of Contrastive Rhetoric (Connor et al., 2008; Connor & others, 1996). However, instead of studying the effect of cultural backgrounds on L2 writing in English (or any other language), we will investigate academic writing traditions in their given lingua-cultural and geopolitical contexts. As such, Contrastive Rhetoric can only be used as a source of inspiration or starting point for identifying relevant features that vary across cultures: the use of metadiscourse (e.g. Ädel, 2008), move structure of texts and sections of texts (e.g. Martín, 2003), reader vs. writer

responsibility (e.g. Kubota, 1997), topical structure of paragraphs (e.g. Almaden, 2008), citation moves in academic journals (e.g. Omizo & Hart-Davidson, 2016) or power distance (e.g. Wolfe, 2008). As more peripheral cultures are likely influenced by the major traditions (see Galtung, 1981), it can be assumed that writing in the Baltic States is influenced by the German and Slavic writing traditions historically and by Anglo-American writing traditions more recently (see Šinkūnienė, 2018). However, it is currently unclear in what respects and to what extent Baltic writing tradition(s) relate to other writing traditions. Nevertheless, our aim is not to compare the Baltic writing traditions with others, but to rely on the contrastive rhetoric approach to provide a methodological basis for identifying and studying features that may vary from writing tradition to writing tradition. As such, our approach will be to regard the Estonian, Latvian and Lithuanian academic writing traditions as having an identity and integrity of their own – collectively or individually – and not simply as collections of features borrowed from other traditions.

As the abovementioned features are mostly manifested above the sentence level, we tap into studies taking a Discourse Analytic approach (Biber, et al., 2007a), especially those that focus on textual features above the sentence level, such as Schiffrin et al. (2001). As we are interested in texts pertaining to specific academic genres, we also draw from previous work done in Genre Analysis (Swales, 1990, 2004), specifically in studies making use of Move and Step analysis (see, e.g., Connor et al., 2007) used to map the rhetorical structure of texts in various genres such as fundraising letters (e.g. Biber, et al., 2007b), patent applications (Groom & Grieve, 2019) or academic texts, such as research articles (e.g. Kanoksilapatham, 2007). Combining different approaches allows us to map the rhetorical structures of academic texts as manifested on macro-, meso- and micro-levels (Fetzer, 2013). By investigating the macro-structures, we determine the general structure of the texts (Lin & Evans, 2012). The meso-level is investigated at section and paragraph levels through move and step structure (e.g. Cotos et al., 2016, 2017) and paragraph structure (Simpson, 2000). Micro-level analysis entails, for instance, how the moves and steps are manifested in lexico-grammatical patterns or by investigating the lexical and grammatical structure of specific concepts: self-promotion, argumentation and other topics relevant for mapping the rhetorical structures in academic discourse.

As indicated earlier, the project provides a methodological advancement (using combined computational, quantitative, and qualitative methods) for rhetorical analysis. In particular, the project aims to break new ground by making use of Machine Learning and Deep Learning methods to process large amounts of authentic academic texts. So far, large-scale computational approaches have hardly been the norm in Contrastive Rhetoric, but this advancement has recently been advocated as an integrated approach to analysing text with other methods, such as corpus-based empirical analysis (Geisler, 2016). Biber et al. (2007a), on the other hand, have emphasized the importance of corpus-based empirical analysis in Discourse Analysis, which traditionally has also deployed other methods to investigate structural organization of texts. In our project, we will compile a large multilingual database of academic texts in Estonian, Latvian, and Lithuanian, which will be collected from available web-based databases and repositories. The genres included will be BA/BSc and MA/MSc and higher-level works, such as doctoral dissertations and journal articles. The data will be collected from University repositories and Open Access journals by web scrapping. Bearing in mind that disciplinary variation can override writing traditions (see, e.g., Martín & León Pérez, 2014; Yakhontova, 2006), we will include texts from across all the different disciplines (humanities, social sciences, natural sciences, medical sciences).

Another innovative aspect of the project will be its combination of both the top-down and bottom-up approaches to analysis. In case of the top-down approach, which prioritises functional frameworks, our units of analysis will be determined before we commence the analysis of our data. In case of the bottom-up approach, which prioritises detailed linguistic description, our units of analysis will emerge during the data analysis process itself (Biber, et al., 2007a). While many authors (e.g. Moreno & Swales, 2018) argue for the top-down approach when it comes to using genre-based move analysis, some studies have deployed a bottom-up approach for similar purposes (e.g. Cortes, 2013). As pointed out by Geisler (2016), there has been some success in using hand coding at a later stage of text analysis, rather than at an early stage (Lejeune, 2011), and in combining automated coding with hand coding at various stages of the text analysis (Lemke et al., 2015). In our project, we will combine the two

approaches to determine the structural patterns in Estonian, Latvian, and Lithuanian academic texts. To achieve this, we will choose and operationalize relevant features, drawing on (but not restricting ourselves to) the "hunches" used in previous literature (Omizo & Hart-Davidson, 2016). The analysis will proceed by developing a coding schema for each feature in all languages. A coding taxonomy will be developed followed by automated analyses using Machine Learning algorithms. In the bottom-up approach, the discourse organization patterns will be detected based on lexico-grammatical patterns (e.g. Biber, Csomay, et al., 2007; Cortes, 2013). The resulting predictive machine learning models can in turn be validated using ethnographic or other qualitative methods (Tse and Hyland 2006; Wiedemann, 2013).

The outcome of the project is thus an empirically grounded model for investigating academic writing traditions. As such, the results are relevant in the Baltic States as the project will create a body of knowledge about the local academic writing traditions. This knowledge can be put to practical use when teaching academic writing in local contexts and when developing teaching materials. It is a well-known fact that, at the moment, there is a lack of original materials for teaching academic writing in Baltic universities and institutes of higher education, and teachers often have to rely on translated textbooks that have been composed for teaching writing in the Anglo-American tradition. The project would be a welcome addition to the few studies in Estonian (Kerge et al., 2014; Meier, 2002; Reinsalu, 2017), Latvian (Dubova, 2019; Dubova et al., 2009; Laiveniece, 2014) and Lithuanian (Gudavičienė, 2019; Linkevičienė & Šinkūnienė, 2012; Šinkūnienė, 2014, 2017, 2018; Volungevičienė, 2018) already available. However, we believe that the results of the project are also relevant globally. The model that we will develop for mapping writing traditions in Estonia, Latvia, and Lithuania can be further used by scholars working in other regions to describe their own writing traditions.

3. Conclusion

To summarise, in this brief position paper, we have demonstrated that there is a knowledge gap not only in terms of what is currently known about more peripheral writing traditions, but also in terms of how these writing traditions might best be studied and described. We have introduced the BWrite

project as an attempt to address both of these issues. The project will contribute to the field of academic writing in many respects: it will create a body of knowledge that is of interest to local communities in the Baltic States and in other regions with uncharted academic writing traditions and it will break new methodological ground in its deployment of Machine Learning and Deep Learning methods for rhetorical analysis. Ultimately, the project aims to offer new and exciting perspectives to scholars working in the fields of Contrastive Rhetoric, Discourse Analysis and Genre Analysis by providing them with a "road map" for detecting in large amounts of data something challenging to detect, namely what is intuitively recognizable as a local writing tradition. Given that the project started in January 2020, the first stage of the project was to present the theoretical framework and foundation for the three languages. The data has been identified and collected, and the three years following will 1) identify, code, and test the features in Estonian, Latvian and Lithuanian (year 2), 2) build models representing the languages and further identify the features across large databases of texts (year 3), 3) identify models, disciplinary differences and test these using more qualitative methods to validate the results (year 4).

Acknowledgment

References

Ädel, A. (2008). Metadiscourse across three varieties of English. *Contrastive Rhetoric: Reaching to Intercultural Rhetoric*, *45*. https://doi.org/10.1075/pbns.169.06ade

Almaden, D. O. (2008). An Analysis of the Topical Structure of Paragraphs Written by Filipino Students. *The Asia-Pacific Education Researcher*, *15*(2), 127–153. https://doi.org/10.3860/taper.v15i2.84

Biber, D., Connor, U., & Upton, T. (2007a). Discourse on the Move. *Using Corpus Analysis to Describe Discourse Structure*. https://doi.org/10.1075/scl.28

Biber, D., Connor, U., & Upton, T. (2007b). Identifying and analyzing rhetorical moves in philanthropic discourse. *Discourse on the Move: Using Corpus Analysis to Describe Discourse Structure*, 43–72. https://doi.org/10.1075/scl.28

Biber, D., Csomay, E., Jones, K., & Keck, C. (2007). Introduction to the identification and analysis of vocabulary-based discourse units. *Discourse on the Move: Using Corpus Analysis to Describe Discourse Structure. Studies in Corpus Linguistics*, *28*, 155–173. https://doi.org/10.1075/scl.28.09cso

Charteris-Black, J., & Ennis, T. (2001). A comparative study of metaphor in Spanish and English financial reporting. *English for Specific Purposes*, *20*(3), 249–266. https://doi.org/10.1016/S0889-4906(00)00009-0

Connor, U., Nagelhout, E., & Rozycki, W. V. (Eds.). (2008). *Contrastive rhetoric: Reaching to intercultural rhetoric*. John Benjamins Pub. https://doi.org/10.1075/pbns.169

Connor, U. & others. (1996). *Contrastive rhetoric: Cross-cultural aspects of second language writing*. Cambridge University Press. https://doi.org/10.1017/CBO9781139524599

Connor, U., Upton, T. A., & Kanoksilapatham, B. (2007). Introduction to move analysis. *Discourse on the Move: Using Corpus Analysis to Describe Discourse Structure*, *2342*. https://doi.org/10.1075/scl.28.04kan

Cortes, V. (2013). The purpose of this study is to: Connecting lexical bundles and moves in research article introductions. *Journal of English for Academic Purposes*, *12*(1), 33–43. https://doi.org/10.1016/j.jeap.2012.11.002

Cotos, E., Huffman, S., & Link, S. (2017). A move/step model for methods sections: Demonstrating Rigour and Credibility. *English for Specific Purposes*, *46*, 90–106. https://doi.org/10.1016/j.esp.2017.01.001

Cotos, E., Link, S., & Huffman, S. (2016). Studying disciplinary corpora to teach the craft of Discussion. *Writing & Pedagogy*, *8*(1), 33–64. https://doi.org/10.1558/wap.v8i1.27661

Dubova, A. (2019). Zinātnes valoda: Konferences materiāli. *Zinātnes valoda : konferences materiāli*, 65. 9789934569661

Dubova, A., Leitāne, M., & Lele-Rozentāle, D. (2009). *Zinātniskā komunikācija starpkultūru kontekstā*. Ventspil Augstskola.

Fetzer, A. (Ed.). (2013). *The Pragmatics of Political Discourse: Explorations across cultures*. John Benjamins Publishing Company. https://doi.org/10.1075/pbns.228

Galtung, J. (1981). Structure, culture, and intellectual style: An essay comparing saxonic, teutonic, gallic and nipponic approaches. *Social Science Information*, *20*(6), 817–856. https://doi.org/10.1177/053901848102000601

Geisler, C. (2016). Current and Emerging Methods in the Rhetorical Analysis of Texts Opening: Toward an Integrated Approach. *Journal of Writing Research*, *7*(3), 417–424. https://doi.org/10.17239/jowr-2016.07.03.05

Groom, N., & Grieve, J. (2019). The evolution of a legal genre: Rhetorical moves in British patent specifications, 1711 to 1860. *Corpus-Based Research on Variation in English Legal Discourse*, 201–234. https://doi.org/10.1075/scl.91.09gro

Gudavičienė, E. (2019). Socialinių ir technologijos mokslų bakalauro darbų kalbos ypatybės. *Žmogus ir žodis*, *20*(1), 4–13. https://doi.org/10.15823/zz.2018.9

Kanoksilapatham, B. (2007). Rhetorical moves in biochemistry research articles. *Discourse on the Move: Using Corpus Analysis to Describe Discourse Structure*, 73–119. https://doi.org/10.1075/scl.28.06kan

Kaplan, R. B. (1966). Cultural thought patterns in inter-cultural education. *1980*, 399–418. https://doi.org/10.1111/j.1467-1770.1966.tb00804.x

Kerge, K., Pajupuu, H., Alp, P., Põlda, H., & Uusen, A. (2014). Towards sophisticated writing. *Proceedings of the Tallinn University Institute of Estonian Language and Culture: Studies in Language Acquisition, Learning, and Corpora*, *16*, 103–115.

Kubota, R. (1997). A Reevaluation of the Uniqueness of Japanese Written Discourse: Implications for Contrastive Rhetoric. *Written Communication*, *14*(4), 460–480. https://doi.org/10.1177/0741088397014004002

Laiveniece, D. (2014). *Zinātniskās rakstīšanas skola*. LiePA.

Lejeune, C. (2011). From Normal Business to Financial Crisis... And Back Again. *Forum: Qualitative Social Research*, *12*(1), 19.

Lemke, M., Niekler, A., Schaal, G. S., & Wiedemann, G. (2015). Content Analysis between Quality and Quantity: Fulfilling Blended-Reading Requirements for the Social Sciences with a Scalable Text Mining Infrastructure. *Datenbank-Spektrum*, *15*(1), 7–14. https://doi.org/10.1007/s13222-014-0174-x

Lin, L., & Evans, S. (2012). Structural patterns in empirical research articles: A cross-disciplinary study. *English for Specific Purposes*, *31*(3), 150–160. https://doi.org/10.1016/j.esp.2011.10.002

Linkevičienė, N., & Šinkūnienė, J. (2012). Asmeniniai įvardžiai mokslo kalboje. *Kalbotyra*, *64*, 78–102.

Martín, P., & León Pérez, I. K. (2014). Convincing peers of the value of one's research: A genre analysis of rhetorical promotion in academic texts. *English for Specific Purposes*, *34*, 1–13. https://doi.org/10.1016/j.esp.2013.09.002

Martín, P. M. (2003). A genre analysis of English and Spanish research paper abstracts in experimental social sciences. *English for Specific Purposes*, *22*(1), 25–43. https://doi.org/10.1016/S0889-4906(01)00033-3

Meier, H. (2002). Olulisi aspekte tekstitüübivõrdluses.- Reet Kasik (toim.), Tekstid ja taustad. *Artikleid Tekstianalüüsist. Tartu Ülikooli Eesti Keele Õppetooli Toimetised*, *23*, 101–114.

Moreno, A. I., & Swales, J. M. (2018). Strengthening move analysis methodology towards bridging the function-form gap. *English for Specific Purposes*, *50*, 40–63. https://doi.org/10.1016/j.esp.2017.11.006

Mur-Dueñas, P., & Šinkūnienė, J. (2016). Self-reference in research articles across Europe and Asia: A review of studies. *Brno Studies in English*, *1*, [71]-92. https://doi.org/10.5817/BSE2016-1-4

Omizo, R., & Hart-Davidson, W. (2016). Finding genre signals in academic writing. *Journal of Writing Research*, *7*(3), 485–509. https://doi.org/10.17239/jowr-2016.07.03.08

Orta, I. V. (2010). A contrastive analysis of the use of modal verbs in the expression of epistemic stance in Business Management research articles in English and Spanish. *Ibérica, Revista de La Asociación Europea de Lenguas Para Fines Específicos*, *19*, 77–95.

Reinsalu, R. (2017). Osutamine iseendale teadustekstis. *Keel Ja Kirjandus*, *11*, 829−845.

Schiffrin, D., Deborah, T., & Hamilton, H. (2001). E. 2001: Introduction. *Deborah Schiffrin, Deborah Tannen & Heidi E. Hamilton (Toim.) The Hand Book of Discourse Analysis*, 1–11.

Severino, C. (1993). The 'Doodles' in Context: Qualifying Claims about Contrastive Rhetoric. *The Writing Center Journal*, *14*, 44–62.

Siepmann, D. (2006). Academic Writing and Culture: An Overview of Differences between English, French and German. *Meta*, *51*(1), 131–150. https://doi.org/10.7202/012998ar

Simpson, J. M. (2000). Topical Structure Analysis of Academic Paragraphs in English and Spanish. *Journal of Second Language Writing*, *9*(3), 293–309. https://doi.org/10.1016/S1060-3743(00)00029-1

Šinkūnienė, J. (2014). Lietuviškojo humanitarinių ir socialinių mokslų diskurso ypatybės. *Podoktorantūros (Post Doc) Stažuočių Įgyvendinimas Lietuvoje: Mokslinių Straipsnių Rinkinys. Vilnius: Baltijos Kopija, 2014*.

Šinkūnienė, J. (2017). Citations in research writing. *Cross-Linguistic Correspondences: From Lexis to Genre*, *191*, 253.

Šinkūnienė, J. (2018). The power of english: I and we in lithuanian, lithuanian english and british english research writing. In *Intercultural perspectives on research writing/edited by Pilar Mur-Dueñas and Jolanta Šinkūnienė* (pp. 59–79). John Benjamins Publishing Company.

Šinkūnienė, J., & Olmen, D. V. (2012). Modal Verbs Of Necessity In Academic English, Dutch And Lithuanian. *Darbai Ir Dienos*, *58*, 153–181.

Swales, J. (1990). *Genre analysis: English in academic and research settings*. Cambridge University Press.

Swales, J. (2004). *Research genres: Explorations and applications*. Cambridge University Press. https://doi.org/10.1017/CBO9781139524827

Tse, P., & Hyland, K. (2006). 'So what is the problem this book addresses?': Interactions in academic book reviews. *Text & Talk*, *26*(6), 767–790. https://doi.org/10.1515/TEXT.2006.031

Volungevičienė, S. (2018). *HUMANITARINIŲ IR TECHNOLOGIJOS MOKSLŲ STUDENTŲ DARBŲ FRAZIŠKUMAS*. 16.

Wiedemann, G. (2013). Computergestützte Analyse qualitativer Daten: Wie sich die qualitative Sozialforschung für Massentextanalysen öffnetOpening Up to Big Data : Computer-Assisted Analysis of Textual Data in Social Sciences. *Historical Social Research Vol. 38*, *No. 4*, Volumes per year: 1. https://doi.org/10.12759/HSR.38.2013.4.332-358

Wolfe, M. L. (2008). Different cultures – Different discourses? Rhetorical patterns of business letters by English and Russian speakers. *Contrastive Rhetoric: Reaching to Intercultural Rhetoric*, *169*, 87. https://doi.org/10.1075/pbns.169.08wol

Yakhontova, T. (2002). Selling'or 'telling'? The issue of cultural variation in research genres. *Academic Discourse*, *10*(1), 216–232.

Yakhontova, T. (2006). Cultural and disciplinary variation in academic discourse: The issue of influencing factors. *Journal of English for Academic Purposes*, *5*(2), 153–167. https://doi.org/10.1016/j.jeap.2006.03.002