

Introduction

Robert Hesselbach , José Calvo Tello ,
Ulrike Henny-Krahmer , Christof Schöch , Daniel Schlör 

1. A New Subdiscipline: *Digital Stylistics*

The increasing access to massive amounts of digital texts in the last decades has allowed different disciplines to emerge or flourish. Disciplines from the humanities (linguistics, literature, history, and philosophy, among others) and computer science (natural language processing, computational linguistics, information retrieval) have created and made use of a new generation of resources, for example corpora, websites, digitized collections, annotation tools, and methods of analysis. Although the disciplines share many interests and methods, there are also specificities about their research objects and the research questions they tackle.

Stylistics is one of the traditional disciplines in the humanities that has benefited from these new resources. The main focus of this discipline is investigating the linguistic style of texts, and literary texts are the object of analysis more frequently than in many other linguistic subdisciplines. In this way, for decades, stylistics has been an interdisciplinary field where linguists and literary scholars have engaged in fruitful discussions.

In the past years, several projects and working groups have coined the concept of *digital stylistics* differently. Digital stylistics combines two previous research fields: digital humanities and stylistics. Digital stylistics is more specific than digital humanities, probably because the researchers needed to use a narrower label to define their community, research objects, and interests since the digital humanities community as a whole has become considerably larger and more diverse in the past decade. However, digital stylistics remains an interdisciplinary field, similar to both digital humanities and stylistics themselves, and it helps to challenge, confirm, or correct existing research paradigms using digital methods.

Moreover, in the past years, several disciplines related to computer science (data mining, machine learning, computational linguistics, artificial intelligence) have influenced and shaped several new subdisciplines in the digital humanities. This influence is visible in many articles in this book, in which methods from computer science are applied to humanities data. The explorative perspective has been widely applied in the past years, for example with the creation of character networks or the use of topic modeling, clustering techniques, or word embeddings. In many cases, this explorative

use is combined with a proper evaluation of the options and parameters that the methods offer. A further possibility is the proposal of new theoretical models, which can be expressed through computational means to structure specific phenomena from the humanities, or the development of new computational methods to investigate specific humanities research questions that could not be answered with previous methods.

Regarding the research object, categories of texts are frequently described by metadata and linguistic features. With categories of texts, we are referring to groups of texts defined, for example, by a common author, the genre, the year of composition or publication, the complexity of the text, or the degree of canonization or literariness. The linguistic features can be extracted from different linguistic levels: phonological, graphemic, morphological, lexical, semantic, syntactic, and pragmatic. In addition, content-related features can be taken into account, for example, the place or time span of the plot. This kind of study of textual categories or groups of texts through linguistic and other types of textual features can be seen in many chapters of this volume. To mention two examples, Laura Hernández-Lorenzo uses lexical information to study two textual categories: authorship and periodization. On his side, Andreas van Cranenburgh looks at specific morphosyntactic units in Dutch to better understand their impact on the perception of literariness.

2. The Founding of *Digital Stylistics* and Its Recent Development

Another evidence for the interdisciplinarity of digital stylistics can be seen in its pioneers, who started analyzing stylistic features with computational means in the 1980s. One of the most frequently cited authors in articles about the style and linguistic particularities of corpora (not necessarily literary ones) is Douglas Biber. His pioneering work has shaped not only the understanding of corpus linguistics but also many other subdisciplines of digital humanities which mainly work with texts. In their article “Exploring Fictional Styles along Universal Dimensions of Register Variation,” Douglas Biber and Jesse Egbert lay out the principles of multi-dimensional analysis and present the results of such an analysis for English across discourse domains. The authors can show how two dimensions, the fundamental opposition between clausal versus phrasal discourse and the opposition between narrative versus non-narrative discourse, play an important role in fictional literature.

On the side of literary texts, John Burrows’ work was an early precursor of the application of computational methods to literature. Digital humanities and, more specifically, the areas of stylometry and digital stylistics are in great debt to this pioneering

scholar. Not only did he employ many methods unknown to the humanities to this point (such as principal component analysis), but he also developed specific new methods himself (Zeta, Delta, Iota) for a range of different research questions.

Since the 1990s, the application of computational and quantitative methods to texts to analyze stylistic features has developed in several directions. A lot of stylistic analysis has been done in corpus linguistics, a subdiscipline that is rooted in linguistics, even though literary texts are also analyzed from time to time. For many years, literary studies did not develop their own subdiscipline for the analysis of corpora, their own *corpus literary studies*, a label that could have been possible but is not a reality in the academic panorama. From the 2000s onwards, several new labels were coined, and one gained momentum during the following years: *digital humanities*, referring to a very broad area of which computational stylistic studies are only one specific part. However, this is not the only label that appears: *corpus stylistics*, *digital stylistics*, and more recently, *computational humanities* and *computational literary studies*. The latter terms focus on the application of computational methods for analysis rather than the representation and presentation of data. The adjective *computational* also emphasizes the importance of quantitative methods. These new labels for subfields are a sign that the very general interdisciplinary field of digital humanities is branching out again.

In order to situate the topic of this publication, *digital stylistics*, in the wider context of different (sub)disciplines related to the digital and the humanities, we ran a series of queries in a library catalog to quantify and visualize its results for the different labels of disciplines. This catalog is called the KIoplus database and constitutes the largest database of library records in Germany. It contains information about monographs (both printed and digital) and journals, but not about chapters, journal papers, or conference contributions.

The queries retrieve the number of publications that contain the labels *humanities computing*, *corpus linguistics*, *digital humanities*, *digital philology*, *corpus stylistics*, *computational humanities*, *computer philology*, *stylometry*, *corpus literary studies*, *cultural analytics*, *digital stylistics*, and *textometry* in their title. To cover the terms in the two most predominant languages of the KIoplus, for the searches, the terms were used in both German and English, and the results for the translations were summed up. The search is defined to also find publications that contain all the tokens of the label in any order. That means a book with the title *A Corpus Study of Literary Works* would be found for the label *corpus literary studies* since the title contains the three words from the label. We also ran exact searches for each label, giving a similar distribution of results but with around a fifth of the number of hits for each label, and many labels with no exact hits.

As Figure 1 shows, *corpus linguistics* and *digital humanities* are the terms that appear most frequently in the titles of the publications, in more than a thousand monographs. *humanities computing* follows these, with more than a hundred monographs.

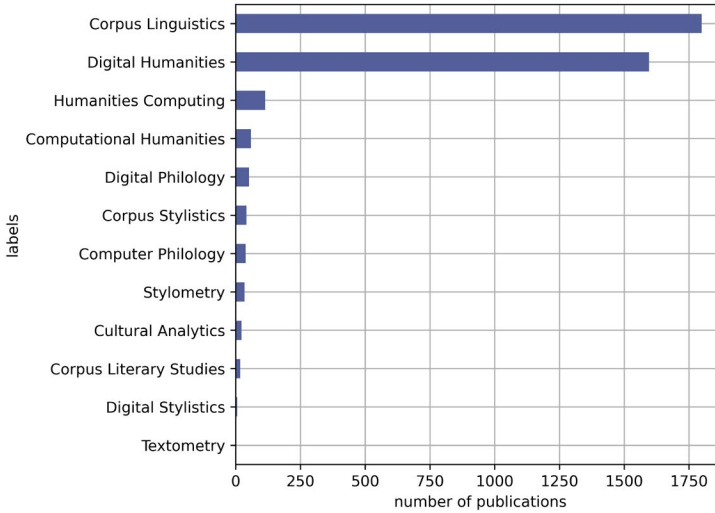


Fig. 1 Frequency of discipline labels in the titles of publications in the database K10plus (Hesselbach, Calvo Tello, Henny-Krahmer, Schöch, Schlör, CC BY).

Next, a group of labels with around forty publications is formed by *digital philology*, *corpus stylistics*, *computational humanities*, and *computer philology*. The following group has around twenty publications populated by labels such as *stylometry*, *corpus literary studies*, and *cultural analytics*. Finally, *digital stylistics* and *textometry* have fewer than ten publications. The results show that the term *digital stylistics*, which is part of the title of the present volume, is not the most widespread one. It addresses a quite specific field of study and community. At the same time, by history, meaning, and research culture, it is related in different ways to all the other terms that were queried and with which the articles published here can be associated, as well, to different degrees.

Given the data presented in Figure 1, one may ask: How has the frequency of these labels developed over time? When did *digital humanities* and *corpus linguistics* gain momentum? Is their frequency still increasing, or have they already reached their peak? To answer this, we ran queries in the same database extracting the frequency of each label from the last 40 years. For better visualization, only the labels with more than 20 publications in total are considered in this second figure.

Figure 2 shows how the number of publications with the label *corpus linguistics* in their title was increasing solidly until the year 2015, with over 100 publications per year. Since then, it seems that the number of titles has been rather stable, with around 90 monographs.

The label *digital humanities* appeared in the late 1990s and showed a steep development in its frequency a decade later. In 2010, only 20 books had this label in their title, but the number increased rapidly during the following years. It appears to peak in the year 2016 with 185 publications and its frequency seems to have stabilized since then with around 160 monographs. The rest of the labels show differences in their chronological frequency. An interesting group of labels is *humanities computing* and *computational humanities*. While the first is one of the most frequent labels up to 2000, the second shows an increase in the number of publications since the 2010s (Biemann et al. 2014), especially since 2020. This return to terms emphasizing the computational in contrast to the digital was predicted by Hockey (2012).

Regarding the labels with fewer publications, *stylometry* shows a similar chronological distribution, with some publications in all analyzed decades. *Corpus stylistics* seems to appear during the decade of 2000, with a few publications each year since then. *Computer philology* starts in the 1990s and keeps appearing in the title of some publications from then on. The labels *corpus stylistics*, *digital philology*, and *corpus literary studies* appear mainly in the twenty-first century. Finally, *cultural analytics* only appeared in the last years of the 2010s. An interesting outcome is that no label appears to have been abandoned completely by the community, as all the analyzed labels keep being used for the titles.

What are the differences between all these subdisciplines? All these labels combine two elements: On the one side, a reference to their main methodological framework:

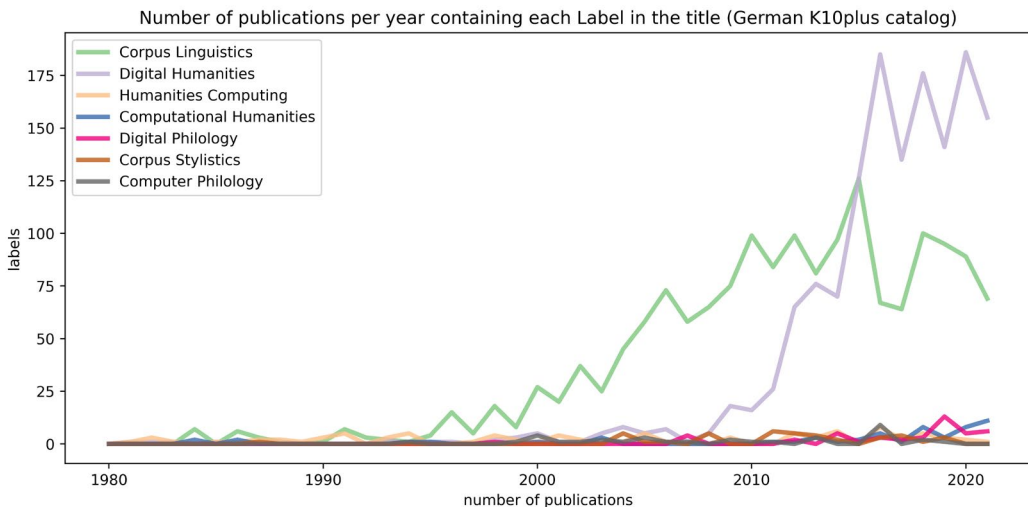


Fig. 2 Frequency of labels in the titles per year in the database K10plus (Hesselbach, Calvo Tello, Henny-Krahmer, Schöch, Schlör, CC BY).

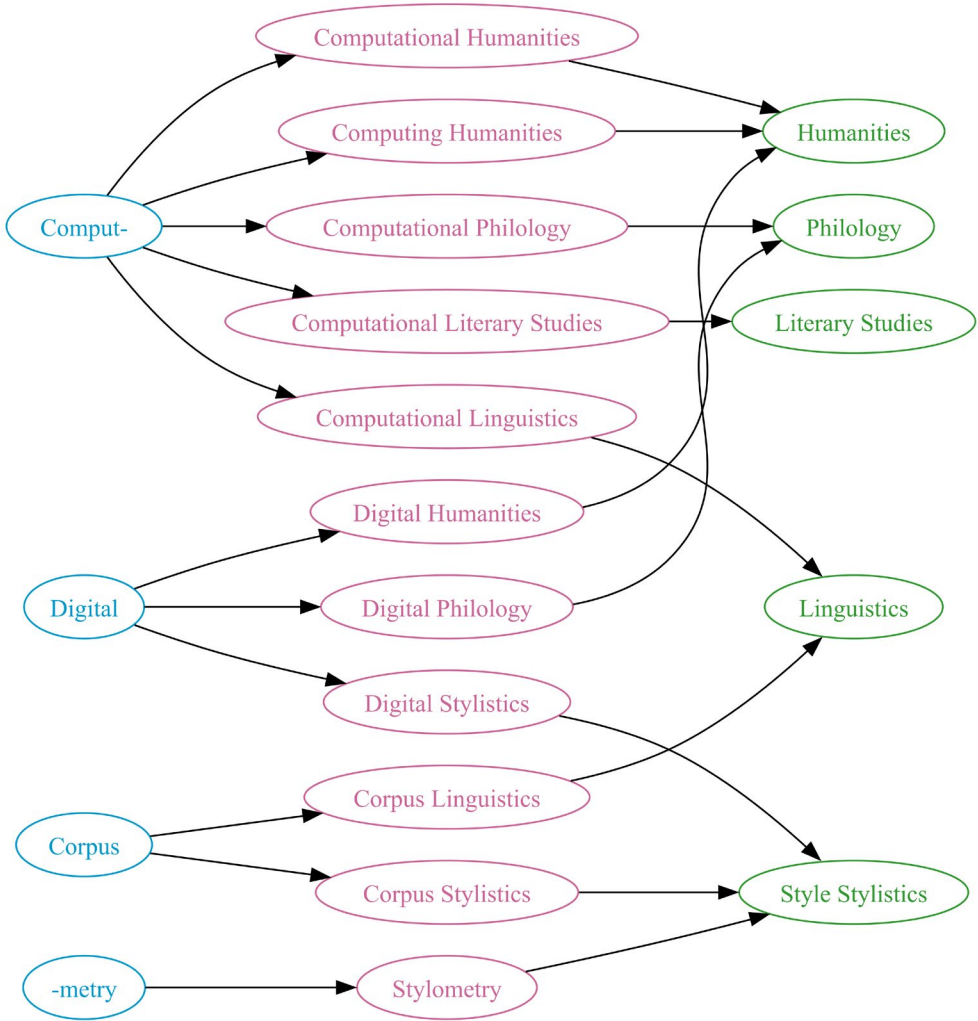


Fig. 3 Combination of the elements of the labels of the subdisciplines (Hesselbach, Calvo Tello, Henny-Krahmer, Schöch, Schlör, CC BY).

corpus, digital, the root *comput-*, or the Greek-based suffix *-metry*. On the other side, there is always a link to a particular discipline from the humanities: linguistics, philology, humanities, literary studies, or stylistics.

By combining both elements, the subdisciplines mark their affinity but also their differences to the other fields. *Computational humanities*, for example, is connected to *digital humanities* and to *computational linguistics*. In that form, the label could be read as a message: our area of study is similar to the digital humanities, however, our methods are more similar to computational linguistics. As this volume shows in its different chapters, digital stylistics shares interests with corpus stylistics and stylometry, applying diverse methods from digital humanities. However, the research field is narrower than that of digital humanities in general. On the other hand, the repertoire of methods differs from the traditional ones in stylometry or corpus stylistics.

We doubt that all these (sub)disciplines and research fields can be understood as mutually exclusive classes. In many cases, they show a great deal of overlap, with scholars publishing or being active in journals, associations, and projects belonging to several of these areas.

3. Digital Stylistics in Romance Studies

Traditionally, the majority of the publications in which literature is analyzed with computational means have focused on English texts. The two pioneers mentioned before, for example, have worked more frequently with English texts than with texts in other languages. Some reasons for this fact are the lack of digitized material, state-of-the-art tools, or financial support for projects concerned with other languages. A further challenge is that, in many cases, the only possible method to analyze multilingual corpora is to first split the texts into monolingual corpora and then run parallel analyses. However, in the last decade, an increasing interest can be noted in working with languages other than English or working with several languages. Some examples of this interest are the versions of the tutorial platform *Programming Historian* in Spanish, French, and Portuguese, the *COST Action Distant Reading* for European Literary History with the *European Literary Text Collection* (ELTeC), the network *Multilingual DH* (<http://multilingualdh.org/en/>) or the special interest group of the Alliance of Digital Humanities Organizations, *Global Outlook::Digital Humanities*. As we will describe later, the *CLiGS* (*Computational Literary Genre Stylistics*) project at the University of Würzburg (Germany) and this volume are part of this increasing interest in analyzing languages other than English.

In the center of this volume are the Romance languages, a group of languages that can be seen as being in the middle range when looking at the available resources.

Often, there are fewer tools and digital linguistic resources, such as dictionaries or annotated data, for these languages than for English. In other cases, available resources have been derived or translated from English, leading to gaps and a cultural bias in them. In many cases, access to large collections of digital texts is more limited in Romance languages than, for example, in German, when resources such as *TextGrid* or the *Deutsches Textarchiv* are considered. However, some Romance languages have a privileged situation in comparison to others, as many articles in this volume do focus on analyzing corpora in Spanish, French, and Italian.

There is also a specificity of the study of Romance languages in the German-speaking area. In other countries the study programs and faculties are split into the national traditions (Spanish Studies, French Studies, etc.), while these are generally grouped together in Romance languages and literatures departments in the German-speaking area. However, in recent years, many digital humanities endeavors have been defined using national or linguistic frontiers. In the last decade, a series of associations have emerged in the Romance languages-speaking countries, such as the *Associazione per l'Informatica Umanistica e la Cultura Digitale*, the *Red de Humanidades Digitales, Humanistica, Humanidades Digitales Hispánicas*, the *Canadian Society for Digital Humanities–Société canadienne des humanités numériques*, just to mention a few from the Americas and Europe.

In the German Romance Studies Association (*Deutscher Romanistikverband*), a working group called Digitale Romanistik on digital humanities in Romance Studies has been active since 2014. During the main conference of Romance studies (*Romanistentag*) in 2021, a specific section dedicated to digital humanities took place for the first time, organized by Digitale Romanistik. Furthermore, in the last year, some interdisciplinary projects with a Romance studies background and a focus on computational methods have been funded, such as the *PhraseoRom* project (Fesenmeier and Novakova 2020), a cooperation between the Université Grenoble Alpes and the universities of Osnabrück, Erlangen-Nürnberg, and Bonn, the above-mentioned project *CLiGS* at the University of Würzburg, or the Mining and Modeling Text project at the Trier Center for Digital Humanities.

4. About this Anthology

The present publication is the outcome of the conference Digital Stylistics in Romance Studies and Beyond, which took place in 2019 at the University of Würzburg from February 27 to March 2. The aim of the conference was to foster research in digital stylistics in the Romance languages and literatures, and to provide an opportunity for international researchers in the field to share their work in order to strengthen the

community of literary scholars and linguists working with these languages. However, the amendment “and Beyond” aimed to keep the conference open for other languages and literatures as well. As digital stylistics has a strong methodological focus and is empirically oriented, the cross-language perspective and exchange of results between different application areas are important to advance the field. This is the reason why in this book one can find contributions on French, Italian, and Spanish as well as other studies on German, English, and Dutch.

The conference was organized by the Early-Career Research Group *Computational Literary Genre Stylistics (CLiGS)*, a project that was based at the Department for Digital Literary Studies at the University of Würzburg and funded by the German Ministry for Education and Research (BMBF) from April 2014 to March 2020. The overall aim of the CLiGS project was to create a convergence between recent methods for the quantitative analysis of literature, on the one hand, and fundamental research questions from genre theory and stylistics, on the other. Several subprojects were realized, focusing primarily on French classical theater as well as Spanish and Spanish-American novels from the nineteenth to the twentieth century (see Schöch 2017; Calvo Tello 2021; Henny-Krahmer 2023). The focus of the *CLiGS* project on Romance literatures and the investigation of literary genres with quantitative and stylistic methods formed the basis for the conference call and it also influenced the kinds of contributions that are published in this volume. All the research results presented in the articles of this book are based on the empirical, computational analysis of literary corpora chosen to analyze and compare either major genres or subgenres of a particular major genre (poetry, drama, prose).

By focusing on genres and subgenres for Spanish, **José Calvo Tello** analyzes in his article “Classification of Genres through 500 Years of Spanish Literature in CORDE” the development of numerous genres over several centuries with the diachronic corpus CORDE (*Corpus diacrónico del Español*) of the Spanish Royal Academy (*Real Academia Española*, RAE). Several classification tests analyze the key factors for each category, finding the length of the text to be a good predictor. A historical analysis furthermore suggests a certain stability of the genres over time.

The next section, devoted to the computational analysis of poetry, begins with a contribution by **Laura Hernández-Lorenzo** (Seville). The author works with Golden Age Spanish poetry and Fernando de Herrera’s poems to analyze Herrera’s role in the stylistic evolution from Renaissance to Baroque and to verify if the posthumous edition of his poetry, *Versos* (1619), is more Baroque, as some critics have suggested. Results point to the transitional role of Herrera’s work in general, with the detection of a more Baroque component in the *Versos* edition. Subsequently, **Nanette Rißler-Pipka** (Göttingen, now Bonn) deals in her contribution “Cross-Language Stylometry: Picasso’s Writings in Spanish and French” with the poetry of one of the most famous Spanish painters, Pablo Picasso. The hypothesis that Picasso’s writings and poetry

are characterized by a unique style is tested with a Spanish and French corpus, and in comparison to contemporary writers, the difference and distinctiveness can be shown by cross-language analyses. After articles on French and Spanish, **Jan Rohden** (Bonn) concludes the poetry chapter with a contribution on Italian. In his article “Digital Approaches to Poetic Style: A Quantitative Stylistic Analysis of Italian Petrarchism,” Rohden raises the question of to what extent digital methods can provide new impulses for research on Petrarchism. A quantitative stylometric analysis of a corpus of Italian love poetry is conducted to identify stylistically distinctive elements of Petrarchism.

The following two chapters focus on the analysis of Spanish and French drama, respectively, and starts with a contribution on the Spanish theater of the Golden Age by **Álvaro Cuéllar** (Vienna, now Barcelona). In his article “Stylometry and Spanish Golden Age Theater: An Evaluation of Authorship Attribution in a Control Group of One Hundred Undisputed Plays,” the author presents a study on 100 Spanish Golden Age theater plays whose authorship is undisputed in order to evaluate which algorithms and which text length are effective for authorship attribution. In his study on French drama (“Repetitive Research: Spitzer and Racine”), **Christof Schöch** (Trier) attempts to retrace Leo Spitzer’s (1887–1960) famous stylistic reading of the tragedies of French seventeenth-century author Jean Racine (1639–1699) using digital text collections and computational methods of analysis, not only revealing new insights into Racine’s and the classical period’s style but also serving to highlight the respective strengths and limitations of established and computational approaches to stylistic analysis.

The following articles all explore narrative literature, with a focus on the nineteenth century at the beginning of this section. In her article “Family Resemblance in Genre Stylistics: A Case Study with Nineteenth-Century Spanish-American Novels,” **Ulrike Henny-Krahmer** (Rostock) applies the concept of family resemblance in a digital genre stylistics analysis of subgenres of nineteenth-century Spanish-American historical and sentimental novels, with a formal implementation of the concept and quantitative evaluation. Besides the use of digital methods as an approach to soft categorization, this analysis shows that the concept of family resemblance itself undergoes change. **Julian Schröter** (Würzburg, now Munich) discusses how procedures of computational and literary genre stylistics can be built and implemented in order to reconstruct the ways in which genres undergo historical change, finding that genre stylistics should, in general, be based on aesthetic interest. In his study on German literature (“Machine Learning as a Measure of the Conceptual Looseness of Disordered Genres: Studies on German *Novellen*”), Schröter outlines the specific historical situation of the German *Novelle* as well as the basis of an aesthetic historiography of this disordered genre. He suggests machine learning tasks be integrated into a psychological framework interpreting accuracy scores as a measure of the semantic looseness of the concept of a specific genre within historical, literary communities.

Turning to twentieth-century prose, **Douglas Biber** (Arizona) and **Jesse Egbert** (Arizona) present an overview of results of multi-dimensional analysis of English across discourse domains. In their article “Exploring Fictional Styles along Universal Dimensions of Register Variation,” the authors show how the two universal dimensions, the fundamental opposition between clausal versus phrasal discourse and the opposition between narrative versus non-narrative discourse, are of great importance in fictional literature. **Andreas van Cranenburgh** (Groningen) analyzes strong and weak pronouns as a stylistic marker of literariness. In his study “Dutch Strong and Weak Pronouns as a Stylistic Marker of Literariness,” he investigates the case of Dutch with a corpus of literary novels and presents quantitative as well as qualitative judgments. The results suggest that style is a prominent factor in the strong/weak pronoun distinction and that a high proportion of strong pronouns is associated with literary prestige and Dutch authorship. **Robert Hesselbach** (Erlangen-Nürnberg) explores in his contribution (“Investigating the Relation between Syntactic Complexity and Subgenre Distinction: A Case Study on Two Contemporary French Authors”) the ways in which the analysis of the syntactic complexity of a narrative text’s sentences can help distinguish between different literary genres for contemporary French novels (1979–2002). Based on an analysis of both qualitative as well as quantitative features, the author argues that syntactic complexity has very little influence on genre distinction and that the degree of syntactic complexity is more likely to appear as an author-related characteristic. The section concludes with a study (“Digital Stylistic Analysis in PhraseoRom: Methodological and Epistemological Issues in a Multidisciplinary Project”) by **Clémence Jacquot** (Montpellier), **Ilaria Vidotto** (Grenoble), and **Laetitia Gonon** (Grenoble). The three authors analyze a large annotated corpus of novels (in French, English, and German) from the twentieth and twenty-first centuries. A stylistic annotation methodology of this corpus is proposed, which links the phraseological analysis of a large literary corpus together with stylistic issues concerning its formal and literary implications, through the concept of the motif.

The last two contributions in this anthology are concerned with Italian and Spanish literature of the twenty-first century, with **Katharina Dziuk Lameira** (Kassel) focusing on modern Spanish novels. In her article “Complexity and Style of Modern Spanish Literary Texts,” the author discusses the question of whether text complexity can be seen as a dimension of authorial style and the relevance of linguistic features for the description of both aspects for modern Spanish literary texts. First analyses show which features and parameters could be suitable for the description of some of the novels analyzed. **Michele A. Cortelazzo** (Padova), **George K. Mikros** (Qatar), and **Arjuna Tuzzi** (Padova) conclude this anthology with their article “Applying General Impostors Method to the Ferrante Case.” By focusing on the contemporary Italian author Elena Ferrante, the authors address Ferrante’s authorship investigation as a verification problem to analyze whether the real author behind Ferrante’s pseudonym is

among the candidates considered in previous studies for a corpus of 150 novels written by 40 authors and a non-literary corpus of 113 texts signed by 14 different entities. In the literary corpus, Starnone emerged as the most likely author of Ferrante's novels; however, Starnone was not the only possible author since in many non-literary texts, Raja, Martone, as well as the E/O publishing house staff and publishers, seem to have made authorial contributions.

As all the articles in this book are concerned with the stylistic analysis of specific linguistic and literary research data, the research results presented here will only be fully transparent and reproducible if the underlying data is accessible. Therefore, the editors of this volume encouraged the authors to publish their research data in an open format, with an open license allowing for re-use and in a specific, identifiable version corresponding to the state that the data had at the time of writing the articles. To facilitate the publication of the research data, a community was created on Zenodo, a service provided by CERN for the long-term archiving and versioning of data, describing it with appropriate metadata and making it citable through digital object identifiers (DOIs) (Nielsen 2013). A Zenodo community serves to bundle different but related data publications. Several articles in this volume are accompanied by data or code records archived on Zenodo. Research data related to the other articles are, in part, offered elsewhere. The volume itself is also published in Open Access form.

Both the conference and the publication would not have been possible without institutional and personal support. We would therefore like to thank first and foremost the German Federal Ministry of Education and Research (*Bundesministerium für Bildung und Forschung*, BMBF) for its generous funding, which made the *CLiGS* project and thus also the Digital Stylistics in Romance Studies and Beyond conference possible in the first place. Furthermore, we have always been supported by Fotis Jannidis, professor for digital literary studies at the University of Würzburg, and his whole team and, therefore, we would like to express our special thanks to Prof. Fotis Jannidis, Dr. Steffen Pielström, and Barbara Gersitz. We would also like to thank our two student assistants, Lea Kunz and Constanze Ludewig, who ensured that the conference ran very smoothly. With this publication, we conclude the *CLiGS* project in a very satisfactory way and hope that the individual contributions will generate further research in the field of digital stylistics.

ORCID®

Robert Hesselbach  <https://orcid.org/0000-0001-9758-8290>
 José Calvo Tello  <https://orcid.org/0000-0002-1129-5604>
 Ulrike Henny-Krahmer  <https://orcid.org/0000-0003-2852-065X>
 Christof Schöch  <https://orcid.org/0000-0002-4557-2753>
 Daniel Schlör  <https://orcid.org/0009-0001-6983-3719>

References

- Biemann, Chris, Gregory R. Crane, Christiane D. Fellbaum, and Alexander Mehler. 2014. “Computational Humanities – Bridging the Gap Between Computer Science and Digital Humanities.” *Dagstuhl Reports* 7: 80–111. <https://drops.dagstuhl.de/opus/volltexte/2014/4792/>.
- Calvo Tello, José. 2021. *The Novel in the Spanish Silver Age: A Digital Analysis of Genre Using Machine Learning*. Bielefeld: transcript. <https://doi.org/10.14361/9783839459256>.
- Fesenmeier, Ludwig, and Iva Novakova, eds. 2020. *Phraséologie et stylistique de la langue littéraire approches interdisciplinaires* [Phraseology and Stylistics of Literary Language: Interdisciplinary Approaches]. Berlin, Bern, Bruxelles: Peter Lang.
- Henny-Krahmer, Ulrike. 2023. *Genre Analysis and Corpus Design: Nineteenth-Century Spanish-American Novels (1830–1910)*. Universität Würzburg. <https://nbn-resolving.org/urn:nbn:de:bvb:20-opus-319992>.
- Hockey, Susan. 2012. “Digital Humanities in the Age of the Internet: Reaching Out to Other Communities.” In *Collaborative Research in the Digital Humanities: A Volume in Honour of Harold Short, on the Occasion of his 65th Birthday and his Retirement, September 2010*, edited by Marilyn Deegan and Willard McCarthy, 81–92. Farnham: Ashgate.
- Nielsen, Lars Holm. 2013. “ZENODO – An Innovative Service for Sharing All Research Outputs.” Presented at the Joint OpenAIRE/LIEBER Workshop, Ghent, Belgium, May 28, 2013. <https://doi.org/10.5281/zenodo.6815>.
- Schöch, Christof. 2017. “Topic Modeling Genre: An Exploration of French Classical and Enlightenment Drama.” *Digital Humanities Quarterly* 11 (2). <http://www.digitalhumanities.org/dhq/vol/11/2/000291/000291.html>.