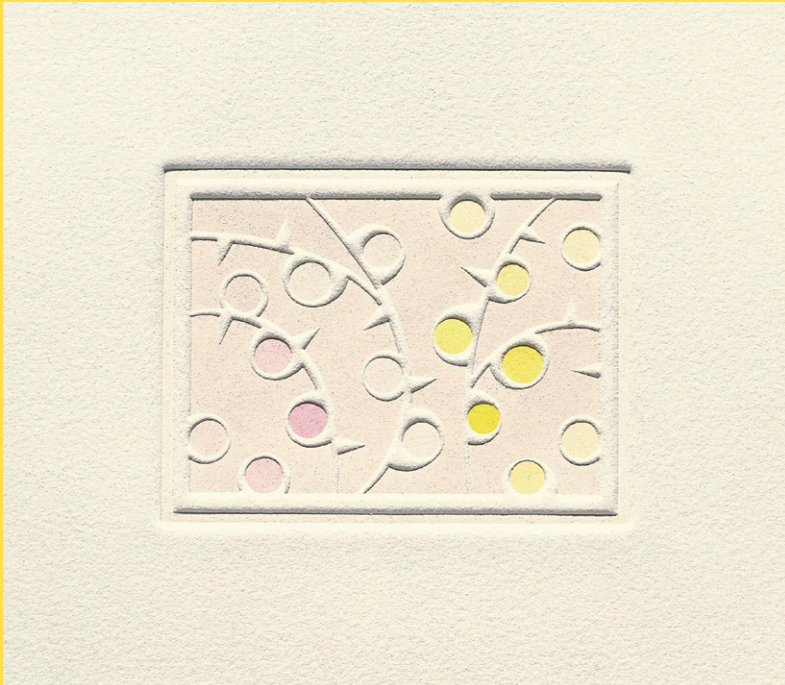


**E**nglish  
**L**anguage  
**O**verseas  
**P**erspectives and  
**E**nquiries



Vol. 21, No. 1 (2024)

ENGLISH STUDIES AT THE DAWN OF THE 21ST CENTURY

Editors of ELOPE Vol. 21, No. 1:  
Bernarda LEVA, Simon ZUPAN and Nastja PRAJNČ KACIJAN

Journal Editors: Smiljana KOMAR and Mojca KREVEL

University of Ljubljana Press  
*Založba Univerze v Ljubljani*

Ljubljana, 2024  
ISSN 1581-8918



**E**nglish  
**L**anguage  
**O**verseas  
**P**erspectives and  
**E**nquiries

Vol. 21, No. 1 (2024)

ENGLISH STUDIES AT THE DAWN  
OF THE 21ST CENTURY

Editors of ELOPE Vol. 21, No. 1:  
Bernarda LEVA, Simon ZUPAN and Nastja PRAJNČ KACIJAN

Journal Editors: Smiljana KOMAR and Mojca KREVEL



University of Ljubljana Press  
*Založba Univerze v Ljubljani*

Ljubljana, 2024

CIP - Kataložni zapis o publikaciji  
Narodna in univerzitetna knjižnica, Ljubljana

811.111'27(082)

ENGLISH studies at the dawn of the 21st century / editors of  
ELOPE vol. 21, no. 1, Bernarda Leva, Simon Zupan and Nastja  
Prajnč Kacijan. - Ljubljana : University of Ljubljana Press = Založba  
Univerze, 2024. - (ELOPE : English language overseas perspectives  
and enquiries, ISSN 1581-8918 ; vol. 21, no. 1)

ISBN 978-961-297-361-2  
COBISS.SI-ID 199658499

# Contents

## PART I: INTRODUCTION

### **Bernarda Leva, Simon Zupan and Nastja Prajnc Kacijan** **9**

English Studies at the Dawn of the 21st Century

## PART II: LANGUAGE

### **Katja Plemenitaš** **15**

“Green, How I Want You Green”: The Discourse of the Brand *I Feel Slovenia*

“Zeleno, ki te ljubim zeleno”: diskurz znamke *I feel Slovenia*

### **Aleksandr Gaisov, Igor Rižnar** **31**

Discourse Analysis of the Mission Statements of Slovenian Enterprises

Diskurzna analiza poslanstev slovenskih podjetij

### **Sándor Czeglédi** **45**

Attitudes of the Hungarian-American Diaspora to the Officialization of English in the United States

Odnos madžarsko-ameriške diaspore do angleščine kot uradnega jezika v Združenih državah Amerike

### **Alexey Tymbay** **63**

On Non-Native Listeners’ Ability to Identify Prominence and Pitch Accents in English Monologic Speech

Sposobnost nerojenih poslušalcev angleščine prepoznavati naglase v angleškem monološkem govoru

## PART III: ENGLISH LANGUAGE AND LITERATURE TEACHING

**Nikola Dobrić**

**91**

Assessment-Relevant Stimuli and Judging of Writing Performances –  
From Micro-Judgments to Macro-Judgments

Dražljaji, pomembni za ocenjevanje, in presoja pisnih izdelkov –  
od mikropresoj do makropresoj

**Tania Tagle, Paulo Etchegaray, Claudio Díaz, Mabel Ortiz,  
Marcela Quintana, Lucía Ramos**

**111**

How Do In-Service EFL Teachers Assess Student Language Learning? Analysis of  
English Assessment Instruments Used in Chilean Secondary Schools

Kako učitelji ATJ ocenjujejo učenje jezika? Analiza instrumentov za ocenjevanje  
znanja angleščine, ki se uporabljajo v čilskih srednjih šolah

**Lena Tica, Ivana Krsmanović**

**129**

Overcoming the Writer's Block? Exploring Students' Motivation and Perspectives  
on Using ChatGPT as a Writing Assistance Tool in ESP

Kako premagati krč pri pisanju? Motivacija in stališča študentov o uporabi  
ChatGPT kot pripomočka pri pisanju v angleščini za posebne namene

**Vesna Petrović**

**151**

Investigating Oral Presentation Assessment in LSP: Practices, Demands  
and Challenges

Ocenjevanje govornih nastopov pri poučevanju jezika stroke: izkušnje, zahteve  
in izzivi

## PART IV: TRANSLATION AND CONTRASTIVE STUDIES

**Katja Dobrić Basanež**

**169**

A Corpus-Based Contrastive Analysis of Metaphorical Collocations of the Terms  
*Right* and *Pravo*: Do Different Legal Cultures Give Rise to Different Metaphors?

Korpusna kontrastivna analiza metaforičnih kolokacij izrazov *right* in *pravo*: ali  
različne pravne kulture ustvarjajo različne metafore?

**Silvana Orel Kos**

**185**

Introduction of Machine Translation into Audiovisual Translation Teaching

Uvajanje strojnega prevajanja v poučevanje podnaslovnega prevajanja

PART V: LITERATURE

**Michelle Gadpaille** 211

The Discourse of the “Good Death” in Nineteenth-Century Fiction and a Historical Case Study

Diskurz »dobre smrti« v pripovedni prozi 19. stoletja in zgodovinska študija primera

**Andrej Zavrl** 229

Authorship Attribution and the Material Realities of Early Modern Play Texts

Ugotavljanje avtorstva in materialnost zgodnjenovoveških dramskih besedil

**Ifeta Čirić-Fazlija** 247

“Slings and Arrows of Outrageous Fortune”: Post-Pandemic Anglophone Theatre and Drama

“Puščic prš proži nezaslišana usoda”: anglofonsko gledališče in dramatika po koncu pandemije

**Vesna Tripković-Samardžić** 265

Contemporary Anglophone Plays on Montenegrin Stages in the Post-Pandemic Era

Sodobna anglofonska dramska dela na črnogorskih odrih v obdobju po pandemiji

**Mahasen Badra** 283

Metatheatre in Brian Friel’s *The Loves of Cass McGuire*: The Semiotics of Make-Believe

Metagledališkost v Frielovi drami *The Loves of Cass McGuire*: semiotika pretvarjanja

**Martina Kastnerová** 299

Archiater Caesarius: Johannes Crato as Philip Sidney’s Forgotten Mentor?

Archiater Caesarius: Johannes Crato kot pozabljeni mentor Philipa Sidneyja?

LIST OF CONTRIBUTORS 312



# **Part I**

## **Introduction**



**Bernarda Leva, Simon Zupan and Nastja Prajnc Kacijan**  
University of Maribor, Slovenia

## English Studies at the Dawn of the 21st Century

In recent years, English Studies has been affected by several global events and developments, with the most notable being the COVID-19 pandemic. Although it has somewhat receded into the past by now, it is worth reminding ourselves that not that long ago, in 2020, the pandemic disrupted the education of 1.6 billion students worldwide, including 220 million tertiary-level students, some of whom were also future language professionals (UNICEF 2021, n.p.). Although many students returned to school within a year, the then World Bank Global Director for Education Jaime Saavedra estimated that “COVID-19 related school closures [were] likely to increase learning poverty to as much as 63%” (UNICEF 2021, n.p.). It need not be highlighted that English was among the disciplines and languages that were affected the most. One testimony to that is a collection of essays with the telling title *The Challenges and Opportunities of Teaching English Worldwide in the COVID-19 Pandemic* (Kılıçkaya, Kic-Drgas, and Nahlen 2022). As the volume demonstrates based on essays by educators from 15 countries on different continents, the pandemic forced teachers to turn their practices upside down and move away from standard routines: they had to modify their teaching and assessment methods; adapt the course content and materials; deal with technical challenges; overcome the digital divide; and find new ways of motivating their students and dealing with student absenteeism, to mention but a few of the challenges they faced. The volume also shows that similar problems were experienced at all levels of education: from pre-university and university institutions to language schools. At the same time, the pandemic proved the resourcefulness of English teachers and students alike (e.g., Moorhouse and Kohnke 2021; Lukas and Yunus 2021; Farkhani, Badiei, and Rostami 2022).

Another prime example of the impact of the pandemic on English Studies is theatre and drama. Following the safety precautions that were put in place, many theatres had to close almost literally overnight and sometimes it took months for them to reopen, and even then they had to operate under restrictions. On Broadway, for example, 31 shows had to close in March 2020 alone, which was unprecedented for an industry that is estimated to support over 96,000 jobs and contribute \$14.7 billion to the New York City economy every year (Solomon 2021). If these effects are extrapolated onto the global level, their implications become even clearer. Fortunately, theatre and drama proved their resilience by quickly adapting and finding an alternative venue online. This applied to the theater-makers, who started producing and streaming shows online; the theatregoers, who watched these productions; as well as the playwrights who began thematizing and reflecting on the pandemic in their plays. As observed by Fuchs, “a great deal of creativity has been unleashed under the extraordinarily difficult conditions of lockdown” (2023, 183). The extent of that creativity is yet to be fathomed.

In addition, recent years have witnessed an unprecedented rate of technology development. Although digitalization has been an integral part of the humanities for over two decades, we

have yet to see its full implications for the future of English Studies. What we have in mind is not just the applicative facets of digitalization such as corpus studies, remote teaching, or “digital theatre”, whose extensions by now are relatively familiar. Instead, it will be much more intriguing to see where artificial intelligence (AI) will take us. Although AI, too, is not new, most of us have only recently become aware of its applications, the best known of which probably is ChatGPT, which has already had a noticeable impact on English Studies. For those of us who are in academia, it has become customary to ask ourselves questions such as the following: How much of the work generated by our students is based on AI? How does one recognize students’ original input? How do we distinguish between the two? Which AI tools are important for us and for the students? How do we incorporate AI in literary studies, in linguistics, in teacher training, or in translation? How do we incorporate AI into teaching, learning, and research? What will be the long-term consequences of AI on language professions in general?

On the other hand, the advent of new technologies does not mean the end of “traditional” English Studies. On the contrary, a random look at the special issues of *Elope* alone in recent years proves that most scholars still explore traditional avenues regarding the English language (e.g., Lipovšek and Ilc 2023; Lauersdorf and Kavalir 2022; Fabijanić and Stopar 2019), literature and culture (e.g., Gadpaille and Mohar 2022; Gadpaille and Blake 2020), language teaching (e.g., Dagarin Fojkar 2019; Hempkin, Kukovec, and Težak 2017) and translation (e.g., Hirci, Pisanski Peterlin, and Zupan 2021). These issues also prove that what typically is *new* are new topics, research methodologies, and tools for examining them, whereas the very core of scholarly interest in English Studies remains relatively stable, to the relief of many.

And this is also the intersection where the present issue of *Elope* is placed: the contributions in it follow perennial themes, addressing them from new perspectives. In the language section, Katja Plemenitaš analyses Slovenia’s tourism brand “I Feel Slovenia”, focusing on the English website’s use of evaluative language. It highlights the term “green” as a value-laden placeholder, integrating various aspects of Slovenia’s image into a cohesive concept. In an earlier corpus-based study, the author also examines the use of the term “green” in the context of environmental changes (Stramljič Breznik and Plemenitaš 2023). Aleksandr Gaisov and Igor Rižnar examine Slovenian MSMEs and large companies’ mission statements across six industries, revealing conservative, company-focused approaches with less stakeholder emphasis. Despite individual variations, they identify common industry-specific themes using similarity and discourse analysis. Sándor Czeglédi presents Hungarian-American communities’ views on federal English officialization, revealing they maintain Hungarian privately but embrace English publicly, opposing Hispanic demands for minority-language accommodations. The section is rounded off by Alexey Tymbay, whose study explains the occasional failure of phonologically trained non-native speakers of English to identify certain accent types: the reason can be found in the listeners’ mother tongues’ prosodic interference in English prominence perception. The ELT section opens with a paper by Nikola Dobrić, who demonstrates that rating writing involves overall impressions rather than discrete elements, yet preconscious stimuli influence these judgments, combining micro-judgments into a complex score. Tagle et al. analyse Chilean EFL teachers’ assessment instruments, revealing a preference for language tests over performance evaluations, emphasizing writing, reading, grammar, and vocabulary. In a paper

about AI in education, Lena Tica and Ivana Krsmanović look at the correlation between English learning motivation and ChatGPT usage, highlighting its practical appeal and mixed user attitudes. The last paper in this section is by Vesna Petrović, who examines LSP teachers' assessment challenges for oral presentations, recommending a unique rating scale to improve evaluation practices.

In the section dedicated to translation and contrastive studies, Katja Dobrić Basanež analyses metaphorical collocations of “right” and “pravo” in English and Croatian legal contexts, revealing cognitive similarities and cultural differences. Silvana Orel Kos, on the other hand, looks at the subtitling trends in Slovenia amid the growth in global audiovisual content and advances in machine translation, analysing student performance in post-editing machine-translated subtitles. In the first paper of the extensive literature division, Michelle Gadpaille explores 19th-century Protestant attitudes toward deathbed scenes, comparing real and fictional narratives to highlight ritualized reassurance and cultural significance. Andrej Zavrl writes about early modern drama's authorship and collaboration through digital analysis, emphasizing challenges in interpreting textual transmission and attribution reliability. Ifeta Čirić-Fazlija's paper examines how theatre reflects societal contexts, noting the COVID-19 pandemic's impact on new genres and themes in British and American drama, exploring whether authors continue to incorporate pandemic experiences into their works. In an earlier study, Čirić-Fazlija (2022) also addresses the manner in which the US theatre has responded to the outbreak of the COVID-19 pandemic. In another paper on post-pandemic theatre, Vesna Tripković-Samardžić analyses the reception and impact of nine contemporary Anglophone plays in Montenegro amidst societal challenges, with insights into theatre's role in addressing global themes and local norms. In another paper on drama, Mahasen Badra explores metatheatrical techniques in Brian Friel's *The Loves of Cass McGuire*, highlighting play-within-a-play, expressionism, and self-reflexive elements akin to Brecht and Pirandello. The final paper in the volume is by Martina Kastnerová, who writes about Philip Sidney's European travels (1572–1575), which profoundly influenced his political and literary growth. This paper explores his interactions with Central European scholars, focusing on Johannes Crato von Krafftheim's influence on Sidney's views on irenicism, medicine, and botany.

As editors of this issue, we want to thank the journal editors Dr Smiljana Komar and Dr Mojca Krevl for accommodating the proposal for this special issue; all the contributors for their papers and for bearing with us and the editorial process; and several dozen anonymous reviewers, who also deserve credit for the volume's scholarly merits. Finally, we must thank the *Elope* Technical Editor, Dr Andrej Stopar, whose continuous support and advice that often went beyond the mere technicalities was instrumental in the making of this issue.

## References

- Čirić-Fazlija, Ifeta. 2022. “Staging race and gender in the era of contemporary crises: Dramas of African American women playwrights” *American, British and Canadian Studies* 39 (1): 54–78. <https://doi.org/10.2478/abcsj-2022-0017>.
- Dagarin Fojkar, Mateja. 2019. “Journeys in Language Education.” *ELOPE: English Language Overseas Perspectives and Enquiries* 16 (2): 9–11. <https://doi.org/10.4312/elope.16.2.9-11>.

- Fabijanić, Ivo, and Andrej Stopar. 2019. "The sounds of English: Introduction." *ELOPE: English Language Overseas Perspectives and Enquiries* 16 (1): 9–11. <https://doi.org/10.4312/elope.16.1.9-11>.
- Farkhani, Zahra Akbarzade, Ghazal Badiei, and Farzad Rostami. 2022. "Investigating the teacher's perceptions of classroom management and teaching self-efficacy during Covid-19 pandemic in the online EFL courses." *Asian-Pacific Journal of Second and Foreign Language Education* 7 (1): 25. <https://doi.org/10.1186/s40862-022-00152-7>.
- Fuchs, Barbara. 2023. *Theater of Lockdown: Digital and Distanced Performance in a Time of Pandemic*. London: Methuen Drama.
- Gadpaille, Michelle, and Jason Blake. 2020. "Introduction: Atwood at 80." *ELOPE: English Language Overseas Perspectives and Enquiries* 17 (1): 9–11. <https://doi.org/10.4312/elope.17.1.9-11>.
- Gadpaille, Michelle, and Tjaša Mohar. 2022. "Celebrating the Precise, the Paradoxical and the 'Pret-ty-Trick-y' in Alice Munro's Fiction." *ELOPE: English Language Overseas Perspectives and Enquiries* 19 (1): 9–12. <https://doi.org/10.4312/elope.19.1.9-12>.
- Hempkin, Kirsten, Melita Kukovec, and Katja Težak. 2017. "Introduction." *ELOPE: English Language Overseas Perspectives and Enquiries* 14 (1): 7–9. <https://doi.org/10.4312/elope.14.1.7-9>.
- Hirci, Nataša, Agnes Pisanski Peterlin, and Simon Zupan. 2021. "Translating in theory and action: contemporary contexts in translation." *ELOPE: English Language Overseas Perspectives and Enquiries* 18 (1): 9–12. <https://doi.org/10.4312/elope.18.1.9-12>.
- Kılıckaya, Ferit, Joanna Kic-Drgas, and Rachel Nahlen, eds. 2022. *The Challenges and Opportunities of Teaching English Worldwide in the Covid-19 Pandemic*. Newcastle-upon-Tyne: Cambridge Scholars Publishing.
- Lauersdorf, Mark Richard, and Monika Kavalir. 2022. "English in Central Europe." *ELOPE: English Language Overseas Perspectives and Enquiries* 19 (2): 9–11. <https://doi.org/10.4312/elope.19.2.9-11>.
- Lipovšek, Frančiška, and Gašper Ilc. 2023. "Novel challenges for discourse analysis: Cross-linguistic and cross-cultural perspectives." *ELOPE: English Language Overseas Perspectives and Enquiries* 20 (1): 9–12. <https://doi.org/10.4312/elope.20.1.9-12>.
- Lukas, Brenda Anak, and Melor Md Yunus. 2021. "ESL teachers' challenges in implementing e-learning during COVID-19." *International Journal of Learning, Teaching and Educational Research* 20 (2): 330–48. <https://doi.org/10.26803/IJLTER.20.2.18>.
- Moorhouse, Benjamin Luke, and Lucas Kohnke. 2021. "Responses of the English-language-teaching community to the COVID-19 pandemic." *RELC Journal* 52 (3): 359–78. <https://doi.org/10.1177/00336882211053052>.
- Solomon, Alisa. 2021. "How Covid transformed US theater." *The Nation*, April 20, 2021. <https://www.thenation.com/article/culture/nonprofit-theater-covid-wpa/>.
- Stramljič Breznik, Irena, and Katja Plemenitaš. 2023. "Zeleno, ki te ljubim, zeleno." *Slavia Centralis* 16 (2): 132–41. <https://doi.org/10.18690/scn.16.2.132-141.2023>.
- UNICEF. 2021. "New global tracker to measure pandemic's impact on education worldwide." March 26, 2021. Accessed June 16, 2024. <https://www.unicef.org/press-releases/new-global-tracker-measure-pandemics-impact-education-worldwide>.

# Part II

# Language



## “Green, How I Want You Green”<sup>1</sup>: The Discourse of the Brand *I Feel Slovenia*

### ABSTRACT

The article examines the discourse of Slovenia's official tourism brand *I Feel Slovenia*. It presents a study of the English version of the official website, focusing on the ways in which the text establishes connections between emotions and emotions in the construction of the brand. Special attention is paid to the role of the word *green* in the discourse construction of Slovenia's official image. The study is based on the framework of evaluative language (Martin and White 1995). The analysis shows that the word *green* is primarily used as a value-laden term, functioning as a placeholder for different evaluative meanings. The word *green* ties different strands of Slovenia's image into a holistic concept rooted in the English-language slogan *I Feel Slovenia* and the label *Slovenia Green*.

**Keywords:** I Feel Slovenia, brand, discourse, English, green

### “Zeleno, ki te ljubim zeleno”: diskurz znamke *I feel Slovenia*

#### IZVLEČEK

Članek obravnava diskurz uradne slovenske turistične znamke *I Feel Slovenia*. Predstavlja raziskavo angleške verzije uradne spletne strani, zlasti ugotovitve o tem, kako se v besedilih konceptualno povezujejo čustva in vrednote trajnostnega razvoja. Analiza se osredotoča zlasti na vlogo besede *zelen* v diskurzni konstrukciji uradne znamke Slovenije. Raziskava temelji na teoriji jezika vrednotenja (Martin in White 1995). Analiza besedil kaže, da se beseda *zelen* uporablja predvsem kot izraz vrednotenja. Deluje kot označevalec vrednot in sodb in povezuje različne sklope slovenske identitete v celostni koncept, ki izhaja iz slogana *I Feel Slovenia* in oznake *Slovenia Green*.

**Ključne besede:** I Feel Slovenia, znamka, diskurz, angleščina, zeleno

<sup>1</sup> A quote from Federico Garcia Lorca's *Romance Sonambulo*, translated into English by William Bryant Logan.

# 1 Introduction<sup>2</sup>

In an economically globalized and socially interconnected world, brands have become ubiquitous indicators of modern consumerism, which penetrate almost every aspect of modern life (Kapferer 2004). The composition of brands should thus be treated as an interdisciplinary subject elucidated by various disciplines, such as economics, sociology, anthropology, history, and semiotics. Furthermore, modern brands have a significant virtual component as communication has increasingly moved to the screen with the expansion of the internet.

Brands confer distinctive identities to products, services, and destinations through verbal and visual communication. According to Yuen (2021, 50), a brand “gives people assurance and peace of mind, as it stands for quality and implies satisfaction with the product”. Building trust is thus one of the crucial elements of building a brand. In an economically globalized and socially interconnected world of the 21st century, it is not just businesses that have to compete for the consumer’s attention in an overcrowded market, but also places and destinations. A particular type of branding developed to increase the identifiability of destinations and even whole countries is called place branding (Ashworth 2009).

This article presents a study of the English-language version of the official brand of Slovenia as it is presented through its official website. The brand itself was created in 2007 by Slovene experts in tourism and advertising in cooperation with different groups of Slovene citizens based on contemporary models of place brands developed in advertising (Konečnik Ruzzier 2012). The primary purpose of the newly created brand, which was to replace the previous less successful attempts at branding Slovenia since its independence, was to construct a contemporary Slovene brand identity for the purposes of promotion, business, and tourism (Konečnik Ruzzier 2012). The brand uses the English-language slogan *I Feel Slovenia*, and the English-language label *Slovenia Green*. The material for the present study includes the wordings of the slogan and the label and selected texts in English accessible from the website *I Feel Slovenia* (Slovenian Tourist Board 2022; <https://www.slovenia.info/en>).

The main purpose of the study is to examine the discourse construction of the image that Slovenia wants to project to the world to attract international visitors. Special attention is given to the interplay of ideational and interpersonal textual elements that characterize the English discourse revolving around the crucial linguistic elements that construe Slovenia’s contemporary brand identity, such as the use of the word *green*.

## 2 The Process of Branding

The theory of branding offers a variety of definitions and classifications of brands and has yet to propose a universally accepted definition of a brand. According to the classic definition

---

<sup>2</sup> This paper was written as part of the ZELEEN.KOM – Komuniciranje podnebne krize za uspešen prehod v zeleno družbo (Communicating the Climate Crisis for a Successful Transition to a Green Society) project, funded by the Slovenian Ministry of Higher Education, Science and Innovation under the Pametna, trajnostna in vključujoča rast; Načrta za okrevanje in odpornost v obdobju 2022–25 programme (Smart, Sustainable and Inclusive Growth Programme; Recovery and Resilience Plans for the Period 2022–25).

by the American Marketing Association (1960), a brand consists of a name, term, design, symbol, or a combination of these intended to identify the goods or services of a business and differentiate them from competitors. Brands foreground the distinctive features of businesses and emphasize what makes them different from their competitors. They are complex symbols that combine various ideas and invoke a body of mental associations that a product has acquired over time (Gardner and Levy 1955). Among features that can make a business distinctive, there are several aspects of a brand, such as a brand as a logo, an image in the consumer's mind, and a value system (de Chernatony and Dall'Olmo Riley 1998). Brand names thus represent the sum of the mental associations they evoke in people (Brown 1992), providing memory shortcuts to consumers under time constraints (Keller 2003). The identity structure of a brand thus includes elements of culture, personality, and self-projection (Kapferer 1992). The net result of a brand is the public image in which the character and personality may be more critical for the overall status of the brand than technical facts about the product (Maurya and Mishra 2012).

Contemporary brands can be classified into different types, such as functional, aspirational, lifestyle, and responsible brands. Functional brands provide functional benefits by fulfilling concrete needs, aspirational brands make us emotionally invested by representing our "dream" products, lifestyle brands help us project an image that is consistent with our lifestyle philosophy, and responsible brands foreground ethical values over the pursuit of profit (Yuen 2021, 39–41). What is notable is that in addition to consumption-related needs, brands fulfil symbolic, internally generated needs for self-enhancement and identity, as well as experiential needs of cognitive stimulation, sensory pleasure, and variety (Park, Jaworski, and MacInnis 1986). Today, the insights gained from the branding process for products have also been applied to cities, countries, and regions. Place branding presents a place or a destination to a variety of potential consumers (domestic and foreign tourists, residents, and investors), shaping their expectations about a place and offering a promise of fulfilment of their various needs. Additionally, a place brand can be used to envision and provide a plan for the future development of a place (Ashworth 2009).

There are several examples of countries that have successfully reinvented themselves by creating a brand to improve their status and reputation, such as Spain, Ireland, and Croatia (Moilanen and Rainisto 2009). These examples show that a successful place branding process can transform a place's image and identity in order to construct an image that engenders positive expectations, associations, and perceptions in potential visitors or investors.

### 3 Slovenia as a Brand

Slovenia's current official tourist brand is *I Feel Slovenia*. According to the Slovene experts who helped create the brand, *I Feel Slovenia* was the first systematically designed and managed brand that presented the identity of Slovenia to the world since its independence in 1991, following several previous, failed attempts (Konečnik Ruzzier and de Chernatony 2013). Konečnik Ruzzier (2012) notes that Slovenia failed to receive the expected recognition from abroad in the initial post-independence period. The problem was that the initial attempts at building Slovenia's brand were reduced to advertising and superficial slogans and logos, rather than a

holistic brand identity. Even more problematic, these slogans and logos frequently changed, further confusing both Slovenes and foreigners and thus not achieving the desired effects.

A place brand aims to distinguish the focal place from other destinations, which can be effectively achieved through stories (Morgan and Pritchard 2002). Previous attempts at promoting Slovenia were missing the consensus on the unique characteristics of Slovenia and the stories that a distinctive brand identity might engender, so the experts designed a questionnaire for focus groups of Slovenes from different walks of life, i.e., stakeholders, to determine their opinions about the distinctive features of Slovenia (Konečnik Ruzzier and Petek 2012).

Based on the analysis of the questionnaire, the authors of the study were able to make data-driven recommendations for the creation of Slovenia's new brand identity. The answers showed that pristine nature, living in harmony with nature, safety, and cultural and natural diversity were considered the critical components of Slovenia's identity. At the same time, the vision for Slovenia's future focused on organic and sustainable development through advanced technology and a niche economy. When asked which colour they most associate with Slovenia, most respondents referred to the colour green as the most distinctive colour of the nation (Konečnik Ruzzier and Petek 2012).

For the new brand, experts created the slogan *I Feel Slovenia* and labels such as *Slovenia Green*, which evoke the values of living in harmony with nature and the pleasant emotions associated with such a lifestyle (Konečnik Ruzzier and de Chernatony 2013). The colour green is used as the shorthand for presenting Slovenia as an attractive, natural, healthy, and safe tourist destination. The centrality of the word *green* for Slovenia's brand is explored in more detail in the analysis presented in Section 5.

## 4 Brands and the Language of Advertising

The language used in brands is linguistically similar to that used in advertising, which is a persuasive type of discourse. Thus it shares similarities with political and argumentative discourse.<sup>3</sup> Classic approaches to advertising are based on semiotic analysis and emphasize the interconnections between verbal and visual modalities and their coded and uncoded elements (e.g., Barthes 1957). According to Crystal and Davy (1983, 222), advertising has two main functions: informing and persuading, with information being subservient to persuasion (Packard 1981). Goddard (1998, 109) points to the grammatical and lexical characteristics of persuasive discourse in advertising, such as the use of comparative references, the connotations of words, "problem-solution" discourse, the use of questions as hook-lines, and playing off written language against its spoken equivalent, such as the use of puns and wordplay. Goddard also mentions the importance of vocabulary connected with the product's unique selling proposition, or "buzzwords" (Goddard 1998, 109). Buzzwords, such as *sustainable* or *environmentally friendly*, convey values and emotions, grab attention, and evoke positive associations in the target audience.

---

<sup>3</sup> The language of persuasion is also frequently used for propaganda purposes in political mobilization (e.g., Hazemali and Onič 2023).

All these principles are taken into consideration in the construction of brands. In creating new brands, special attention is devoted to creating brand names. Brand names “are created to convey information, carry desirable connotations and generate favourable perceptions” (Room 1994, 2). Among word-formation processes, compounding and blending are the most common for creating brand names (Danilović Jeremić and Josijević 2019, 2–3). Homographic overlaps represent another linguistic approach that is commonly exploited for a promotional purpose. A notable example of such creativity can be found in the formation of the slogan for Slovenia’s brand, which uses the homographic overlap between the word *love* and the segment *-love-* contained in the word *Slovenia* (cf. Section 6).

In order to express appropriate values and emotions, advertisers must be familiar with their target audience’s values and expectations and the values they want to convey. According to Hoey (1994), written discourse creates a relationship between the writer and the imagined reader, within which the writer makes assumptions about the ideal reader’s expectations and knowledge. In the persuasive language of advertising, the lexically expressed values suggest assumptions about the ideal customer. In the language of advertising, these assumptions are explicitly or implicitly expressed through attitudinal lexis expressing social values and emotions (Martin and White 2005). Križan (2016, 200), for example, applies the discourse semantic appraisal model by Martin and White (2005) to the analysis of the language of advertising, claiming that this model “is a useful analytical tool for revealing how the language of advertising evaluates the participants it engages, and further, how it helps shape, reflect and naturalize values, norms, and relationship through attitudinal judgment”. The main aspects of Martin and White’s model of evaluative language used in this study are presented in Section 5.

Martin and White’s appraisal framework is used to examine how the word *green* acquires evaluative dimensions through its figurative use and thus plays a central role in the construction of the brand. As noted by Vidaković Erdeljić and Milić (2023), the underlying motivation for using the word *green* figuratively in the context of environmental topics is clearly metonymic, as it involves the activation of a part for whole, whereby the colour term stands for nature in general. The metonymic nature of the word *green* has recently received increased scholarly attention in the context of studies of communication in the time of climate crisis (e.g., Stramljić Breznik and Plemenitaš 2023).

## 5 Materials and Methodology

The present analysis focuses on attitudinal meanings expressed on the English-language version of the website of Slovenia’s official brand.<sup>4</sup> Particular attention is given to capturing the semantics and the role of the word *green* in its evaluative dimension, primarily as an expression of positive appreciation of a variety of values and positive emotions.

The text for the analysis was taken from the English language version<sup>5</sup> of the website of the brand *I Feel Slovenia* (Slovenian Tourist Board 2022; <https://www.slovenia.info/sl>). As

<sup>4</sup> The analysis is applied to the version of the website from March 2022. It has to be noted that some aspects of the website presented in the study no longer exist in the same form.

<sup>5</sup> The languages, in addition to Slovene and English, are German, Italian, French, Russian, Spanish, and Chinese.

mentioned in Section 2, *I Feel Slovenia* is the brand that was developed by Slovene experts based on a systematic survey of different groups of the general Slovene population. The brand also functions as a slogan, forming the logo's tagline. The Slovene Tourist Board manages the website, while the Government Communication Office supervises the brand.

The website of Slovenia's brand is called *I Feel Slovenia*, which is the main slogan of the brand. The logo with the slogan is displayed in the upper-left-hand corner of the website. The slogan is written against a green background (Figure 1). The colour green functions as a recurrent visual motif on the website through its use in graphic signs and its prominence in photographic material. The banner on the homepage contains a section entitled "Explore", which leads to a selection of texts contained on the website. In the version of the website from March 2022, the "Explore" section leads to texts about the Tour of Slovenia, a national cycling competition similar to the Tour de France. The lower part of the banner contains three horizontally organized headlines ("things to do", "places to go", and "plan your trip") which can be clicked to reach further thematically organized sections, ultimately linking to individual articles and videos on the website. Below the banner, there is a short text followed by several sections that are organized vertically and represent another route to reach the articles and videos on the website. They include sections "Stories from Slovenia", "Inspiring places", "Slovenia at a glance", "Taste Slovenia", "Developing green Slovenia", "My way of planning new adventures", "My way of finding new means of relaxation", and "My way of discovering the local character". Each headline covers three or four clickable photographs with taglines that lead to the articles on the website. Towards the bottom of the website there is a section reserved for social media content, such as Instagram, Facebook and Twitter, information on travel and safety guidelines, and a business section.

A website can be regarded as a macro-genre, a genre within which it is possible to identify a range of other genres (e.g., Martin 1992). Texts posted on a website are usually connected through links and function as hypertext. The *I Feel Slovenia* website resembles a network structured hypertext, with a starting point in the sections on the home page but no particular ending point. The links attached to every text enable multiple paths and can lead back to the texts on the homepage in a cyclical fashion. The network structure of the website thus encourages digressions. It keeps the readers engaged, prompting them to open new texts and thus get immersed in more and more promotional texts about Slovenia and its tourist offers. The links to the texts are grouped into several sections on the home page, and most of these links are presented as a picture with a tagline. The analysis includes texts that are directly linked from the home page of the website, a total of 26 texts, in addition to headlines and captions. The examined textual material consists of 10,868 words. The analysis of the visual material is outside the scope of the study due to space limitations. However, other sources provide a deeper insight into the inter-modal connection between text and pictures of the Slovenia brand (e.g., Ivanuša 2021).

The readers can access the website in eight different languages.<sup>6</sup> The present analysis focuses on the English text version of the website. The brand name (and the slogan) exist only in English, and there is no official Slovene version of the name or slogan.

---

<sup>6</sup> The languages, in addition to Slovene and English, are German, Italian, French, Russian, Spanish, and Chinese.

The study presents a qualitative analysis of the use of the word *green* in selected texts from the official Slovene tourist website, including the slogan and labels. For the purpose of capturing the full semantic nature of the word *green* combined with various headwords, the analysis adopts some aspects of Martin and White's categorization of evaluative lexis in the system of interpersonal meaning called appraisal. The appraisal model is a systemic-functional model of evaluative language used to categorize interpersonal meanings that are construed through lexis and grammar. In Martin and White's appraisal theory (2005), primary attention is given to the linguistic resources of the basic appraisal categories of graduation, engagement, and attitude. Engagement is concerned with the way the value position is aligned between the author and the audience, i.e., whether the value position is presented as one which can be taken for granted for a particular audience or as one which is in some way contentious (Martin and White 2005, 92), while graduation deals with linguistic resources for upgrading or downgrading evaluative meanings (Martin and White 2005, 135). Attitude, however, deals with expressions of positive and negative evaluations expressed explicitly or implicitly as lexical items.

The present analysis is mainly concerned with attitude, which subsumes the lexical expression of personal and moral judgments of character (judgment), emotion (affect), and evaluations of things (appreciation). Judgment is the evaluation of human behaviour expressed through the categories of normality (how special one is, e.g., *lucky*), tenacity (how dependable one is, e.g., *brave*, *headstrong*), capability (how capable one is, e.g., *clever*, *stupid*), veracity (how honest one is, e.g., *frank*, *disloyal*) and propriety (how far beyond reproach one is, e.g., *kind*, *rude*) (Martin and White 2005, 52). Affect is an expression of emotions, which are divided into the categories of (dis)inclination, e.g., *love*, *hate*, (in)security, e.g., *safe*, *unsafe*, (un)happiness, e.g., *happy*, *sad*, and (dis)satisfaction, e.g., *content*, *discontent* (Martin and White 2005, 45). Finally, appreciation is concerned with the evaluation of things. This category, in particular, is essential for the expression of values, as it is attributed to objects or abstract referents rather than people. According to Martin and White (2005, 56), appreciation can be divided into 'reactions' to things (if they catch our attention; if they please us, e.g., *beautiful*, *captivating*, *ugly*), their 'composition' (balance and complexity, e.g., *intricate*, *simple*), and their 'value' (how innovative, authentic, timely, significant, e.g., *groundbreaking*, *innovative*, *sustainable*).

## 6 Findings and Discussion

### 6.1 The Construction of the Slogan and Labels as Frames of Slovenia's Brand

The colour green dominates in the visual representation of Slovenia's brand through its use in the design of the logo and other graphic signs, and the display of photographic material with a focus on the green countryside. At the same time, the importance of the colour *green* is also emphasized through discourse, as it is used as an attribute with a wide variety of referents, including the whole country of Slovenia, as illustrated by one of the headlines on the website, "Developing green Slovenia".

The importance of the colour green for Slovenia's brand is explicitly acknowledged in various sections of the website, most notably in its business section that explains Slovenia's brand:

Slovenia has a brand with a meaningful message – I feel **Slovenia**. The Slovenia brand combines all areas of Slovenia, including tourism, and a mix of emotions, sensibility,

and the Slovenian green colour are at its core and identity. (Slovenian Tourist Board 2022; <https://www.slovenia.info/en/business/slovenian-tourist-board/i-feel-slovenia>)

Notably, the colour green is referenced through the nominal phrase “the Slovenian green colour”, in which the adjective *Slovenian* classifies the colour based on national identity. This identity-based reference to green is further framed by the attitudinal meaning of affect, as the phrase “the Slovenian green colour” is directly coordinated with expressions of positive affect (“a mix of emotions and sensibility and the Slovenian green colour”), all defined as the elements at the core of Slovenia’s identity. The “Slovenian green” is further described as “more than a colour” and as an expression of “a balance between the calmness of nature and the diligence of the Slovenian people” (Slovenian Tourist Board 2022; <https://www.slovenia.info/en/business/slovenian-tourist-board/i-feel-slovenia>). With this additional framing through the attitudinal meanings of positive appreciation (“the calmness of nature”) and positive judgment (“the diligence of the Slovenian people”), the word *green* becomes semantically imbued with all three categories of attitudinal meaning used in its immediate context, i.e., affect, judgment, and appreciation (cf. Martin and White 2005).

The evaluative meaning of positive affect plays a significant role in the discourse of Slovenia’s brand. It is also expressed in the main slogan, *I Feel Slovenia*, through the highlighted segment of the slogan. The logo with the slogan, which is displayed on the website, consists of a non-symmetrical trapezoid with a green background containing the line *I Feel Slovenia* (Figure 1). The line is written inside the trapezoid in capital letters. The words *I, feel*, and the string of letters coinciding with the segment *-love-* in the word *Slovenia* are marked white, while the remaining segments of *Slovenia*, i.e., *-s-* and *-nia*, are marked with a pale shade of green. The highlighted segments form the sentence *I Feel Love*. To achieve this evaluative effect, the language of the slogan exploits the homographic overlap between the English word *love* and a four-letter segment *-love-* from the word *Slovenia*, reaffirming positive emotions associated with Slovenia.



FIGURE 1. The official logo of the brand *I Feel Slovenia* (Slovenian Tourist Board 2022; website *I Feel Slovenia*; <https://www.slovenia.info>).

In addition to the slogan, another important framing element is the certificate *Slovenia Green* (Figure 2). *Slovenia Green* is the main certificate conferred to destinations and accommodations that fulfil the Tourist Board’s criteria of being environmentally friendly.<sup>7</sup>

<sup>7</sup> The brand also contains the certificate *Green & Safe*, which was of particular importance during the Covid pandemic.

The graphic sign for the *Slovenia Green* certificate consists of a non-symmetrical trapezoid with a light background and a leafy green-grey design on the right side (Figure 2). The segment marked in dark green highlights *Love Green*. Similarly to the slogan *I Feel Slovenia*, it uses the homographic segment *-love-* inside the word *Slovenia* to achieve an evaluative framing of positive affect. The marked segment can have different syntactic interpretations: it can be read as an imperative sentence, i.e., ‘love green!’, as an elliptical sentence, i.e., ‘I/we love green’, and perhaps more unusually, as an adjectival or nominal phrase with *green* as the headword and *love* as a nominal premodifier (i.e., ‘green of the type associated with love’). Furthermore, the full phrase *Slovenia Green* can also have multiple syntactic interpretations: it can be read as an adjectival or nominal phrase with the adjective *green* as the headword and the noun *Slovenia* as the premodifier (i.e., ‘green of the Slovene type’). Similarly, the line can be interpreted as a nominal phrase, with the word *green* as a noun. Perhaps less probably, it can also be read as an elliptical sentence with a missing linking verb (i.e., ‘Slovenia /is/ green’).



FIGURE 2. The official label *Slovenia Green*: the example *Slovenia Green Destination Platinum* (Slovenian Tourist Board 2022; website *I Feel Slovenia*; <https://www.slovenia.info/sl/poslovne-strani/zelena-shema-slovenskega-turizma/druzina-slovenia-green>).

Both the slogan *I Feel Slovenia* and the certificate *Slovenia Green* express the evaluative meaning of positive affect.<sup>8</sup> The wording of the slogan *I Feel Slovenia* achieves this in the following way: the word *Slovenia* functions in a syntactic position which is typically filled by a noun expressing an emotion or sensation. Due to its syntactic positioning, the word *Slovenia* acquires the semantic flavouring of emotion. Second, the highlighting of the segment *-love-* in the word *Slovenia* in the same line reinforces the conceptual connection between Slovenia and the emotion of love, i.e., feeling Slovenia is akin to feeling love, which also leads to happiness. According to the appraisal theory, love is a subcategory of happiness (cf. Martin and White 2005, 50).

The wording of the certificate *Slovenia Green*, on the other hand, establishes a special interpretation of the colour green based on the Slovene national identity through the premodification of the word *green* by the noun *Slovenia*. Like the slogan, it exploits the segment *-love-* in the word *Slovenia*, showing the word *green* in the syntactic position of the object of affection (i.e., *love green*). With this sleight of hand, it manages to create a conceptual

<sup>8</sup> The analysis of evaluative meaning in the wordings of the slogan and the certificate, including the highlighted segments, is applied to their conventional syntactic readings.

tie between Slovenia, the feeling of love, and the colour green. The explicit explanation of the elements of the brand further strengthens the connection between Slovenia and positive affect.

The wording *Slovenia Green* also semantically opens the headword *green* to different evaluative interpretations based on the category of appreciation, depending on which values associated with Slovenia are focused on in a particular context. Consequently, the certificate of *Slovenia Green* can be attributed to a wide variety of actors that align with the value systems of the brand. The main values expressed by the label are concerned with the areas of environmental protection and lifestyle, and in particular the preservation of a healthy environment and the benefits of living in harmony with nature. *Slovenia Green* thus functions as an attribute with a wide range of different referents, including *Slovenia Green* destinations, providers, accommodations, *Slovenia Green* parks, travels, agencies, and even attractions.

The general framing elements of the brand, i.e., the slogan, certificates, and the official explanation of the brand, however, are all connected through the expression of love as a category of positive affect.

It is worth noting that the wordings of the slogan and certificates, such as *I Feel Slovenia* and *Slovenia Green*, only exist in the English language version. Even the plaques awarded to the Slovene tourism providers lack a Slovene translation of the slogan, labels, or accompanying text.<sup>9</sup> Martina Drakulić, a Slovene expert on contemporary tourism trends, is critical of the lack of official Slovene translations (personal correspondence), but notes that Slovene tourism providers have accepted English-language slogans, certificates, labels, and English borrowings of new terms in tourism (e.g., *nesting*, *glamping*). In the field of tourism, the use of English language slogans and certificates for Slovene providers is often justified by the fact that their primary target audience consists of international visitors, who account for 80% of all tourist visitors (Drakulić, personal correspondence).

Figure 3 shows the plaque with the certificate *Slovenia Green* awarded to Bohinj. The plaque exists exclusively in its English-language version.

Throughout the website, the colour green is strongly associated with a variety of feelings belonging to the category of positive affect, including feelings of trust and safety, in addition to love. This is further shown in Section 6.2, which examines the evaluative dimensions of the word *green* in selected texts from the website.

## 6.2 Evaluative Dimensions of the Word *Green* in the Main Sections of the Website

In this section, the evaluative dimensions of the word *green* are examined in the discourse of the texts contained in the main sections displayed on the home page of the website.

On top of the webpage, the banner prominently displays four evaluative terms expressed as a series of sentential fragments, i.e., “Nature. Freedom. Passion. My Way”. The word *nature* is

---

<sup>9</sup> The author is indebted to Jana Vilman for drawing attention to the lack of Slovene translations of the awarded plaques (personal correspondence).



FIGURE 3. The plaque with the English-language certificate *Slovenia Green* awarded to Bohinj (source: personal archive by Jana Vilman).

followed by three different evaluative expressions: the word *passion*, which has the meaning of positive affect, the word *freedom*, which has the meaning of the positive judgment of capability, and the phrase *my way*, which contains an implicitly expressed positive judgment of determination (cf. Martin and White 2005). The word *nature* is thus strongly associated with evaluative meanings. This kind of evaluative framing is also common with the word *green* used in the discourse of the website.

The headline “Explore” inside the banner leads to an article about the Tour of Slovenia, an international cycling event. The tour is described as “more than just a cycling race” – it is a “journey across a green country”. This wording echoes the description used in the definition of the brand (cf. Section 6.1), where the colour green is described as being “more than just green”. The association of the race with the colour green in the phrase “a race across a green country”, confers a special significance to the events and activities of the tour. This section offers additional meta-messaging on the double meaning of the colour green, as the term *green* is explicitly acknowledged as a term with both a literal and figurative meaning. Green is the colour of “the green jersey” of the tour’s overall winner, but also of Slovenia’s movement “towards sustainable development”.

The text featured below the banner is entitled “Slovenia is waiting for you to explore it. In your way”. In this text, the word *green* is associated with the feeling of safety and isolation from the rest of the world. The text describes Slovenia as a “small region offering so much”, “a perfect green and safe oasis”, and “the green heart of Europe”. The metaphors entailed in the use of the nouns *oasis* and *heart* highlight the smallness, isolation, centrality, and uniqueness, i.e., an oasis is a small, flourishing place surrounded by a desert, and a heart is the small, central organ that propels the body. As their attribute, the word *green* acquires an evaluative meaning of positive appreciation, i.e., of the value of smallness, coupled with the value of diversity (“a small region offering so much”) and exclusivity (“boutique experiences”). In addition to the category of positive affect, i.e., feeling safe and positive appreciation as exclusivity, the text also foregrounds positive judgment of the hosts’ behaviour as responsible and positive appreciation with regard to the experiences offered by Slovenia as enjoyable and unforgettable.

The next section, entitled “Stories from Slovenia”, offers further strengthening of the evaluative dimension of the word *green*. The introduction to a promotional film entitled “A documentary that reveals the green orientation of Slovenia” focuses on the importance of the colour green for Slovenia, talking about Slovenia as a “green destination” and as a country with a “green orientation” having “green ideas”. The word *green* premodifies abstract referents, i.e., the country’s goals and ambitions, and thus transforms into a future-oriented concept denoting telos and tenacity in referencing the aspiration and determination of a whole country. The metonymical use of the word *green* is also explicitly acknowledged: “[Y]ou will learn that it [‘Slovenia’] is truly green. Literally, in the colours of its endless forests and figuratively, in the effort of the locals for it to remain so”. The use of the intensifier “truly” in the phrase “truly green” shows that the adjective *green* is used as a gradable evaluative concept. The figurative level of the word *green* is explained as Slovenia’s efforts to preserve nature and develop sustainably. These efforts are also implicitly connected to the positive judgment of the capabilities and character attributed to Slovenes (“the locals”), who are presented as hardworking and responsible. In addition to being an adjective, the word *green* is used adverbially (e.g., “A documentary that makes you think only green”), denoting a way of thinking. The adjective *green* is also used as an attribute of the noun “story”, i.e., a semiotic entity, in the phrase “a green story of Slovenia,” which conceptually summarizes all the different uses of the word *green*.

The section entitled “Inspiring places” uses the word *green* predominantly as an attribute of place nouns. This use is framed by all three categories of attitudinal meaning: positive appreciation, expressed as a category of aesthetic evaluation and sustainability, is combined with positive judgment, expressed as a category of capability, i.e., health, and positive affect as a feeling of well-being and self-confidence. The Slovene capital, Ljubljana, for example, is described as “a lively green city combining the charm of a small capital and the self-confidence of large European cities” and as “the green capital of a green country”. Here, the value of smallness is combined with positive affect, i.e., “self-confidence”. Well-being as a category of positive affect is strongly associated with health as the positive judgment of capability, both of which are attainable in “every corner of green Slovenia”, where you can find “natural healing remedies, each of which has its own beneficial effect on health and well-being”. Framed by the category of positive affect, in particular well-being and relaxation, the colour *green*

is also used as an attribute of a variety of abstract entities denoting activities, feelings, and experiences related to holidays, which are all framed by positive affect, i.e., the feelings of well-being, relaxation and taking a break, e.g., “green relaxation at a thermal spa”, “pamper yourself in a green way”, and “a green break at Mala Nedelja Bioterme”.

The section “Discovering green Slovenia” towards the bottom of the website discusses Slovenia’s policy with regard to the development of tourism by mentioning the Green Scheme of Slovenian Tourism. The use of the adjective *green* in the name of the policy reinforces the link between the colour green and the values of sustainability and responsibility. The article states that under the auspices of the Green Scheme of Slovenian Tourism, “tourism is developed in a responsible and sustainable way”, through the promotion of “green policies” and “green experiences”. The adjective *green* thus additionally functions as an attribute to entities with institutional meaning, such as *policy* and *scheme*. Slovenia prides itself on being one of the first countries to adopt a green certification scheme, which has been implemented since 2015. The certification scheme reinforces the connection between emotions and values through the positive affect of trust, a feeling that comes from shared values and connects visitors, tourist providers, and destinations: “Now it’s time for more trust between us”.

The article at the bottom of the homepage, entitled “Five green and safe reasons for me to choose Slovenia”, presents additional framing of the brand *I Feel Slovenia* in the form of a listicle with five items. The article focuses on Slovenia being a safe, clean country with high hygiene standards, which was of particular importance during the time of the COVID-19 epidemic. The adjective *green* is paired with the adjective *safe*, whereby *green* can be interpreted as a co-hyponym of *safe*, thus both expressing positive appreciation as a value of security. They are used as attributes of the headword *reasons*, i.e. an abstract noun denoting logic, which expands the range of headwords with *green* as a premodifier even further. A prominent display of the sign with the wording *Green & Safe* is accompanied by the exclamation “Welcome to green and safe Slovenia”, repeating the pairing of *green* and *safe* from the headline. This framing adds an important component to the adjective *green*, as it associates it with freedom from disease.

At the end of the section, the *Slovenia Green* label is described as an umbrella term encompassing the different values and feelings the word *green* is imbued with. Its main purpose is to encourage tourism providers “to go green”. This is another indication that being or becoming *green* is a conscious choice that can be attributed to a wide variety of referents.

### 6.3 The Diversity of the Use of the Word *Green*

In the discourse on the website, the word *green* is predominantly used as an adjective in attributive position with a range of different headwords. Its use in nominal form is rare. Exceptions would be an alternative interpretation of the adjectival phrase *Slovenia Green* as a nominal phrase and the use of the noun *greenery* derived from the base *green*, i.e., “the embrace of greenery at Terme Dobrna”. In one case, *green* is also used as an adverb, indicating a way of thinking, i.e., “A documentary that makes you think only green”. This adverbial use is similar to the use of the adjective *green* in an adjunct realized as a prepositional phrase, i.e., “pamper yourself in a green way”.

The adjective *green* is used mostly figuratively with a range of different headwords. In most cases, the figurative use involves metonymy, with a few examples occurring as part of a metaphor, e.g., a reference to hops used in the brewing industry as “green gold”, or the description of Slovenia as “the green heart of Europe”.

Furthermore, the adjective *green* is rarely used in its literal sense of the colour of vegetation or other features of nature. Even in its literal meaning, it tends to be framed by positive appreciation, such as the value of accessibility to an active life, e.g., “Green nature is only a step away from any city centre”, or aesthetic appreciation, e.g., “wonderful green nature”.

In its metonymical figurative use, the adjective *green* becomes an attribute of a variety of referents, denoting both concrete and abstract entities. They include concrete entities (e.g., “green products”), places (e.g., “green haven”, “green sanctuary”, “green land”, “the green capital of a green country”, “green Slovenia”, and “green hotels”). They also include abstract entities, such as activities and states (e.g., “green experiences” and “green life”), abstract concepts of cognition or logic (e.g., “green ideas”, “green mindset”, “green buzz of sustainability”, “green reasons for choosing Slovenia”). Furthermore, they include semiotic referents (e.g., “a green story”) and institutional entities (e.g., “green policies” and “a green scheme”).

The figurative meanings of the adjective *green* evoke a host of positive evaluative meanings and present a promise of the various experiences that can be enjoyed in Slovenia. The colour green is connected to positive affect in the form of love or comfort, well-being, and relaxation; it is further associated with the positive judgments of Slovene people as reliable, responsible, and healthy, and it also represents broader values in the form of positive appreciation, such as safety, sustainability, progress, uniqueness, diversity, authenticity, and a healthy lifestyle. However, these values are in the service of exclusivity rather than inclusivity, as Slovenia’s striving for “boutique experiences” shows that the green experience is not aimed at the masses. It is also interesting to note that aesthetic evaluations, e.g., “wonderful green nature”, are not foregrounded in the discourse of the brand. This may be a conscious choice of the creators of the brand to avoid sounding too trite and clichéd.

In the process of absorbing evaluative meanings of different categories, the word *green* becomes a placeholder that has the potential to be filled with a wide variety of evaluative content.<sup>10</sup> The physical aspect of the colour green is referenced rarely, mostly when the adjective *green* modifies the headword *nature*. Even when there is a clear focus on colour as a physical category, it is framed by evaluative meanings. This is illustrated by the following sentence: “For years, we have strived to act in a sustainable and responsible manner towards its wonderful green nature and rich local cultural heritage, setting an example for others with our actions”. In this sentence, “green nature” is explicitly framed by several evaluative meanings: positive appreciation as an aesthetic value (i.e., “wonderful green nature”), positive appreciation as a category of sustainability (i.e., “sustainable...manner”), and positive judgment based on moral behaviour (i.e., “responsible.....manner”; “an example for others”).

---

<sup>10</sup> This semantic space, however, is not open to negative figurative meanings of the word *green*, such as the meaning of being immature or envious, e.g., *green with envy*, *green behind the ears*.

Although positive judgments and values feature prominently in the discourse surrounding the brand, it is the category of positive affect, i.e., the emotions of love and trust, that functions as the general evaluative frame, connecting visitors, tourism providers, and destinations by recognizing their shared values: “Now it’s time for more trust between us”.

## 7 Conclusion

Slovenia’s brand, *I Feel Slovenia*, ties together elements of the image that contemporary Slovenia has selected to present itself to the world. The English language has a significant role in conveying the brand, since the slogan and related labels, certificates and plaques are realized only in the English language. By highlighting the segment *-love-* in the word *Slovenia*, the English-language slogan frames the discourse of the brand with positive emotion, establishing a conceptual connection between feeling love and experiencing Slovenia.

Another vital element in the discourse construction of Slovenia is the word *green*, which is always framed by positive evaluations expressed in its context. The analysis of the texts from the main section of the website confirms the use of the word *green* as an evaluative term. The word *green* is used metonymically to modify a wide variety of referents, from concrete entities, such as features of the landscape, places, or products, to abstract referents, such as activities, semiotic entities, and institutions. In this way, the adjective *green* functions as a placeholder for feelings and different positive values and judgments that present the basis of the construction of Slovenia’s brand. With emphasis on the evaluative meanings of love, responsibility, and sustainability, *I Feel Slovenia* presents itself both as a lifestyle brand and as a responsible brand.

Future research could include a wider variety of promotional texts in different languages for corpus-based research into the modern discourse of promotion and tourism, in particular in the context of the digital and green transitions.

## References

- American Marketing Association. 1960. *Marketing Definitions: A Glossary of Marketing Terms*. Chicago: American Marketing Association.
- Ashworth, Gregory J. 2009. “The instruments of place branding: How is it done?” *European Spatial Research and Policy* 16 (1): 9–21. <https://doi.org/10.2478/v10105-009-0001-9>.
- Barthes, Roland. 1957. *Mythologies*. Paris: Editions du Seuil.
- Brown, Gordon. 1992. *People, Brands and Advertising*. Warwick: Millward Brown International.
- Crystal, David, and Derek Davy. 1983. *Investigating English Style*. London: Longman.
- Danilović Jeremić, Jelena, and J. Josijević. 2019. “To blend so as to brand: A study of trademarks and brand names.” *Lexis. Journal in English Lexicology* 14: 1–19. <https://doi.org/10.4000/lexis.3732>.
- de Chernatony, Leslie, and Francesca Dall’Olmo Riley. 1998. “Modeling the components of a brand.” *European Journal of Marketing* 32 (11/12): 1074–90. <https://doi.org/10.1108/03090569810243721>.
- Gardner, Burleigh B., and Sidney J. Levy. 1955. “The product and the brand.” *Harvard Business Review*, March–April, 33–39.
- Goddard, Angela. 1998. *Language of Advertising. Written Texts*. London: Routledge.
- Hazemali, David, and Tomaž Onič. 2023. “‘Canning the Kaiser’ in words and images: Case studies of patriotic American propaganda from WW1.” *ELOPE: English Language Overseas Perspectives and Enquiries* 20 (2): 49–63. <https://doi.org/10.4312/elope.20.2.49-63>.

- Hoey, Michael. 1994. "Signalling in discourse: a functional analysis of a common discourse pattern in written and spoken English." In *Advances in Written Text Analysis*, edited by Malcolm Coulthard, 26–45. London: Routledge.
- Ivanuša, Ema. 2021. "Building a brand: A multimodal analysis of the brand *I Feel Slovenia*." M.A. thesis, University of Maribor.
- Kapferer, Jean-Noël. 1992. *Strategic Brand Management*. London: Kogan Page.
- . 2004. *The New Strategic Brand Management: Creating and Sustaining Brand Equity Long Term*. London: Kogan Page.
- Keller, Kevin Lane. 2003. "Understanding brands, branding, and brand equity." *Interactive Marketing* 5: 7–20. <https://doi.org/10.1057/palgrave.im.4340213>.
- Konečnik Ruzzier, Maja. 2012. "Developing brand identity for Slovenia with opinion leaders." *Baltic Journal of Management* 7 (2): 124–42.
- Konečnik Ruzzier, Maja, and Leslie de Chernatony. 2013. "Developing and applying a place brand identity model: The case of Slovenia." *Journal of Business Research* 66 (1): 45–52. <https://doi.org/10.1016/j.jbusres.2012.05.023>.
- Konečnik Ruzzier, Maja, and Nuša Petek. 2012. "The importance of diverse stakeholders in place branding: The case of 'I Feel Slovenia'." *Anatolia. An International Journal of Tourism and Hospitality Research* 23 (1): 49–60. <https://doi.org/10.1080/13032917.2011.653631>.
- Križan, Agata. 2016. "The language of appraisal in British advertisements: The construal of attitudinal judgement." *ELOPE: English Language Overseas Perspectives and Enquiries* 13 (2): 199–220. <https://doi.org/10.4312/elope.13.2.199-220>.
- Martin, James R. 1992. *English Text: System and Structure*. Philadelphia: John Benjamins Publishing.
- Martin, James R., and Peter R.R. White. 2005. *The Language of Evaluation*. Basingstoke: Palgrave Macmillan.
- Maurya, Upendra Kumar, and Pahrads Mishra. 2012. "What is a brand? A perspective on brand meaning." *European Journal of Business and Management* 4 (3): 122–33.
- Moilanen, Teemu, and Seppo Rainisto. 2009. *How to Brand Nations, Cities and Destinations: A Planning Book for Place Branding*. Basingstoke: Palgrave Macmillan.
- Morgan, Nigel, and Annette Pritchard. 2002. "Contextualizing destination branding." In *Destination Branding: Creating the Unique Destination Proposition*, edited by N. Morgan, A. Pritchard, and R. Pride, 10–41. Oxford: Butterworth-Heinemann.
- Packard, Vance. 1981. *The Hidden Persuaders*. Harmondsworth: Penguin.
- Park, C. Whan, Bernard J. Jaworski, and Deborah J. McInnis. 1986. "Strategic brand concept-image management." *Journal of Marketing* 50 (October): 135–45. <https://doi.org/10.1177/002224298605000401>.
- Room, Adrian. 1994 [1991]. *NTC's Dictionary of Trade Name Origins*. Lincolnwood: NITC Business Books.
- Slovenian Tourist Board. 2022. *I Feel Slovenia*. Accessed June 5, 2022. <https://www.slovenia.info/en>.
- Stramljič Breznik, Irena, and Katja Plemenitaš. 2023. "Zeleno, ki te ljubim zeleno." *Slavia Centralis* 16 (2): 132–41. <https://doi.org/10.18690/scn>.
- Vidaković Erdeljić, Dubravka, and Goran Milić. 2023. "Should green always get the green light? Translation of figurative technical terms in view of conceptual integration theory." *ELOPE: English Language Overseas Perspectives and Enquiries* 20 (2): 49–63. <https://doi.org/10.4312/elope.20.2.9-29>.
- Yuen, Royce. 2021. *Decoding Branding*. London: Routledge.

**Aleksandr Gaisov**

University of Primorska, Slovenia

**Igor Rižnar**

University of Primorska, Slovenia

2024, Vol. 21 (1), 31-43(316)

journals.uni-lj.si/elope

<https://doi.org/10.4312/elope.21.1.31-43>

UDC: 81'42:658(497.4)

## Discourse Analysis of the Mission Statements of Slovenian Enterprises

### ABSTRACT

The study analyzes the discourse of a sample of mission statements of Slovenian Micro, Small and Medium Enterprises (MSMEs) and large companies across six different industries. First, we assess how distinctive these statements are, both at the level of individual companies and within entire industries, using a custom similarity analysis. Then, through a discourse analysis approach, we examine the methods employed by these companies when creating their mission statements. The results reveal that Slovenian organizations often rely on conservative, company-focused mission statements that place less emphasis on stakeholders. Additionally, despite significant variations among individual statements, there are common themes observed when these statements are grouped by industry.

**Keywords:** mission statements, discourse analysis, text similarity analysis

### Diskurzna analiza poslanstev slovenskih podjetij

#### IZVLEČEK

Članek analizira diskurz poslanstev slovenskih podjetij v različnih gospodarskih panogah. Najprej smo s pomočjo analize podobnosti, ki smo jo sami razvili, ocenili stopnjo edinstvenosti poslanstev na ravni posameznih podjetij in po panogah. Zatem smo z uporabo analize diskurza ugotavljali, kako slovenska podjetja oblikujejo diskurz poslanstev. Ugotovili smo, da so poslanstva slovenskih podjetij pogosto precej konservativne narave, saj so osredotočena na podjetje, pri čemer redko omenjajo notranje deležnike. Poleg tega pa so si poslanstva kljub individualnim razlikam med seboj precej podobna, če jih združimo po panogah. V zaključku predlagamo prihodnje raziskave na obravnavanem področju in podamo nekatere omejitve naše raziskave.

**Ključne besede:** poslanstva podjetij, diskurzna analiza, analiza podobnosti besedil

# 1 Introduction

Strategic management, as defined by David (2011, 6), comprises: “integrating management, marketing, finance/accounting, production/operations, research and development, and information systems to achieve organizational success.” This organizational process is split into three stages: development, execution, and evaluation of a strategic plan. The creation of a *mission statement* takes place at the first stage, along with the identification of the *vision* and *values*, SWOT analysis, and the establishment of long-term goals (David 2011, 6).

A corporate mission statement, in turn, is a formal declaration of a company’s core objective, identifying its products and market focus (David 2011, 11). The ongoing deviation from this basic notion greatly contributes to the ambiguity that exists between the concepts of the mission, vision, and values of a company (Salem Khalifa 2011). Some researchers even argue that a single mission statement is enough because it already includes all the elements of vision and values (Powers 2012; Stallworth Williams 2008). Furthermore, due to the lack of standards and common practices, mission statements differ in a variety of ways, including inconsistent length, scope of meaning, and terminology used, among others (Dermol 2013).

Mission statements have gained in popularity in the last three decades (Alegre et al. 2018; Stallworth Williams 2008) mainly due to their perceived positive impact on operational performance, strategic planning, organizational culture, leadership, and communication with stakeholders (Dermol 2013).

Therefore, a key purpose of mission statements is to communicate with a company’s *stakeholders* (David 2011, 49; Klemm, Sanderson, and Luffman 1991), both internal and external, each with different goals and motivations. Internally, mission statements inspire and coordinate employees, while externally they enhance brand recognition and foster external relationships (Alegre et al. 2018). Although these processes are not mutually exclusive, effectively conveying a company’s mission to these distinct groups may require diverse modes of communication. As a result, companies often incorporate their mission statements into various documents, such as annual reports, loan requests, agreements with suppliers and customers, labour contracts, and internal business plans (David 2011, 43).

In this sense, most Slovenian senior managers of medium and large businesses are aware of the primary functions of mission statements as both strategic and managerial tools. Nevertheless, micro and small firms are more reliant on direct communication with their stakeholders (Babnik et al. 2014).

The relationship between mission statements and business performance is a topic of ongoing debate, and the research findings are varied. Researchers approach this link in different ways, some examining individual components within mission statements, while others consider the presence or absence of the entire statement.

Thus, for US and Canadian organizations, researchers generally agree that three major components within mission statements – *purpose*, *values*, and *competitive advantages* – are positively correlated with company performance, particularly regarding employees (Bart and Baetz 1998; Bart and Hupfer 2004; Pearce II and David 1987). In contrast, Slovenian firms

tend to emphasize *the growth of the business, cooperation, and partnership* in their mission statements (Babnik et al. 2014; Breznik 2012; Dermol 2013).

The link between companies' performance and the presence of a mission statement is also ambiguous. Alegre et al. (2018) identified several major groups of studies that have examined various aspects of corporate performance in connection with mission statements, including their impact on employees' behaviour, customer orientation, financial performance (especially return on sales and assets), effects on management, and communication with stakeholders. Furthermore, Dermol (2013) suggests a positive correlation between the presence or absence of a mission statement and the value added per employee, a critical business performance indicator.

Although these and other studies report the positive influence of mission statements on various performance measures (Desmidt, Prinzie, and Decramer 2011), some researchers argue that there is not enough empirical evidence to establish a strong link between mission statements and performance (Stallworth Williams 2008). The evidence typically demonstrates a weak-to-moderate correlation with individual performance measures, and studies tend to focus on a limited set of such measures, making their selection subjective.

In this sense, Salem Khalifa (2011) emphasizes that high-performing organizations might excel in many areas, including creating better mission statements, suggesting that the connection between mission statements and business performance is not straightforward cause and effect. Researchers' conclusions are further undermined by the limitations of their studies, such as sample size and sampling methods.

On this premise, we argue that organizations confront at least two key issues in terms of mission statements. First, they must consider how to effectively communicate their purpose to both internal and external stakeholders. This involves examining the meaning embedded in the company's mission statement and how it is understood and interpreted by individuals within and outside the organization.

Second, assuming that a company's purpose contributes to its prosperity, it is desirable for that purpose to be unique. In other words, a firm cannot effectively compete in the market if it offers similar things to its clients as its competitors do. Therefore, if a company's purpose reflects what sets it apart and its mission statement is aligned with that purpose, it must be special and stand out from other companies' mission statements.

Slovenian companies are no less affected by these challenges and are no exception in their need for high-quality mission statements. Therefore, this study aims to understand and evaluate how Slovenian companies employ mission statements to convey their purpose to stakeholders, and how distinctive these mission statements are at both individual company and industry levels.

One way of approaching this is through *discourse analysis* (DA). DA is a method of studying *discourse*, which comprises meaningful symbolic communication through *language*. This includes both the communicative action itself and its surrounding sociocultural *context*

(Johnstone 2008, 2–3). Thus, analysing mission statements using DA goes beyond just the text and considers the social context, allowing for a better understanding of individual mission statements and the intended message of the author.

DA gained prominence in the 1970s, particularly through the works of the French philosopher Michel Foucault, who is considered a key proponent of discourse theory (Pedersen 2009). Foucault challenged the notion of linear development in discourses, suggesting that they can experience abrupt shifts in previously held beliefs and meanings. These transformations are driven by a combination of internal and external factors and are central to Foucault's principle of discontinuity, which states that every historical period has its own distinctive way of looking at things (Choque Aliaga 2018).

Hence, Foucault emphasized the role of language in shaping ideas, leading to commonly accepted knowledge that gets institutionalized and carried forward in future discourses. In this sense, DA is a way to study the history of knowledge, that is, its creation, evolution, and the limitations imposed by language, institutions, and technology (Pedersen 2009). Therefore, DA focuses on understanding the dynamics of ideas and nodal points of discourse and follows three core directives (McHoul and Grace 2015, 49): discourse must be considered within its historical context, identifying the prerequisites that led to its emergence, and focusing on practical rather than theoretical applications of discourse.

This made DA particularly relevant in areas that study interpersonal interaction, social identity, ethnicity, culture, and language (Johnstone 2008, 7), evolving into an interdisciplinary field associated with anthropology, linguistics, sociology, and psychology (Fairclough 2013a, 4). Due to the internal dynamics and inherent biases of these domains, the study of discourse has branched into two primary models: one based on linguistics and sociolinguistics, the other rooted in sociology and social psychology.

The two models diverge in their primary areas of focus. The linguistic model places its focus on language and its linguistic attributes, including grammar, phonology, morphology, syntax, and semantics, within social contexts. Conversely, the sociological model is more focused on social phenomena, where language is just one component (Taylor 2013, 2–3).

Both approaches emphasize the interconnectedness of discourse, language, and the sociocultural context. This connection comprises the power relations (Fairclough 2013b, 46–51), knowledge development through language usage (Dijk 2014, 5–12), and meaning formation that all occur through various social interactions (Gee 2011, 16). As a result, various justifications for this relationship have led to the development of different methodologies within the field of DA. These methodologies range from formal *linguistic analysis*, *empirical analysis* of conversations and genres, to *critical analysis* of sociocultural elements (Hodges, Kuper, and Reeves 2008).

While there could be some value in examining individual mission statements in detail, our main focus in this study is on a broader perspective. Therefore, we opt for critical discourse analysis (CDA) as our approach. CDA operates within a three-dimensional framework that includes evaluations of: (1) the discourse's products (in our case, mission statements); (2)

how they are created, shared, and used; and (3) the sociocultural context in which they exist (Fairclough 2013a, 59).

## 2 Materials and Methods

We drew a convenience sample from the population of Slovenian companies with domestic capital sources, ensuring they originated from Slovenia. The data was collected from the SloExport website (SloExport 2022), which is based on annual surveys conducted on firms in the Slovenian market, updated in 2022. For each company, we recorded their name, size (micro, small, medium, or large), industry, and mission statement.

Certain companies have explicitly shared their mission statements, clearly indicating them with formatting features like bullet points, bold fonts, or titles within their company descriptions. However, in cases where mission statements were not immediately apparent due to unclear beginnings and endings, we identified them based on the presence of the word “mission” and related terms, as well as their contextual position within the text:

- At the beginning of a sentence: “Company X’s mission is...”, “Our main goal is...”, “Our primary activity is...”, etc.
- In the middle of a sentence: “...with the mission to...”, etc.
- At the end of a sentence: “...is our mission”, etc.

Occasionally, these sentences were surrounded by other sentences that did not exactly fit these criteria, but still expressed the fundamental goals of a company. In such cases we combined them into a single mission statement to ensure we did not miss any relevant data.

TABLE 1. Representation of industries within the sample.

Industry	Number of companies	Percentage of the sample
Agriculture	19	18%
Commerce	24	22%
Electronics, IT and Telecoms	14	13%
Industrial Manufacturing and Materials	21	20%
Services	19	18%
Transport and Logistics	10	9%
TOTAL	107	100%

Subsequently, we determined the six most prevalent industries among the sampled companies, excluding those industries that constituted less than 9% of the sample. The remaining industries, along with their updated respective sizes, are presented in Table 1.

To achieve our primary objective, the method was structured into two empirical sections. The first section featured a *similarity analysis* (SA) of mission statements within the research sample and served a dual purpose. It evaluated the similarity among the mission statements of Slovenian companies, as well as categorizing them by industry to identify more distinct fields. In turn, the SA employed in this study included two complementary techniques: *term frequency–inverse document frequency* (TF–IDF) weighting and *cosine similarity*.

TF-IDF is a statistical approach to calculate the importance of a term based on its frequency within a specific document and across a set of documents (Sheng, Wei, and Yang 2018). In this sense, higher TF-IDF weights indicate greater importance (Qaiser and Ali 2018). In our case, mission statements were treated as documents and individual words as terms. The process was automated using a Python (Python 2022) script and the Scikit-learn (Pedregosa et al. 2011) library's 'TfidfVectorizer' class.

Once the TF-IDF weights for each mission statement were obtained, the subsequent step involved determining their similarities through a cosine similarity matrix, a widely recognized approach for determining document similarity (Manning, Raghavan, and Schütze 2009, 121). This is achieved by comparing the cosine values of two term vectors within the documents. As a result, the cosine similarity value reflects the degree of relevance between the documents (Rahutomo, Kitasuka, and Aritsugi 2012). The Scikit-learn library also provides a 'cosine\_similarity' method that takes the TF-IDF weight matrix as input and computes the cosine similarity for each pair of rows (representing companies' mission statements).

The method returned a symmetric matrix with a size equivalent to the total number of documents in the corpus (i.e., 107 mission statements). Values within this matrix spanned from zero to one, with one signifying identical mission statements and zero indicating a lack of common terms. Based on the cosine similarity matrix, it was deduced that *agriculture*, followed by *electronics*, *IT and telecoms*, along with *transport and logistics*, emerged as the three most distinct industries in terms of their mission statements.

In the second empirical section, Fairclough's (2013b) CDA was applied to the mission statements of three randomly chosen Slovenian companies representing distinct industries: *Pomurske Mlekarne d.d.* (agriculture), *Metra Inženiring d.o.o.* (electronics, IT and telecoms), and *Dars d.d.* (transport and logistics). In turn, CDA is divided into three key stages (Fairclough 2013b, 21–22): *description*, *interpretation*, and *explanation*.

The description stage involved text analysis, which is shaped by productive and interpretative processes. These processes are influenced by participants' individual knowledge, beliefs, stereotypes, and experiences, collectively referred to as the *members' resources* (Fairclough 2013b, 20). Members' resources impact the formal aspects of discourse, such as vocabulary, grammar, and textual structures (Fairclough 2013b, 92–93).

Additionally, Fairclough (2013b, 93) identified three types of values derived from these discourse aspects: *experiential* (agents' knowledge and experiences), *relational* (relationships between agents), and *expressive* (agents' social identity).

During the interpretation stage, we focused on understanding how discourse processes are influenced by members' resources and context (Fairclough 2013b, 117–18). Therefore, we analysed the situational and intertextual contexts to determine discourse type and presuppositions (Fairclough 2013b, 122–27). Interpretation also involved assessing language use, semantics, internal connections, and the main point of the discourse (Fairclough 2013b, 119–20).

In the explanation stage, we explored how discourse interacted with broader social structures. This included examining power relations, ideologies, and the sustainability of the discourse at situational, institutional, and societal levels. Furthermore, it involved studying the utilization of members’ resources in the context of the discourse (Fairclough 2013b, 135–38).

### 3 Results

The results of the SA (the cosine similarity matrices) were visualized as heatmaps via two commonly used Python libraries: Pandas (McKinney 2010) and Seaborn (Waskom 2021). To simplify visualization, mission statements were sorted alphabetically according to their respective industries, allowing for reference through index numbers rather than displaying the full list of company names directly on the heatmaps. The indices for companies’ mission statements, grouped by industry, are thus provided in Table 2.

TABLE 2. Indices of companies’ mission statements grouped by industry.

Industry	Range of indices
Agriculture	0–18
Commerce	19–42
Electronics, IT and Telecoms	43–56
Industrial Manufacturing and Materials	57–77
Services	78–96
Transport and Logistics	97–106

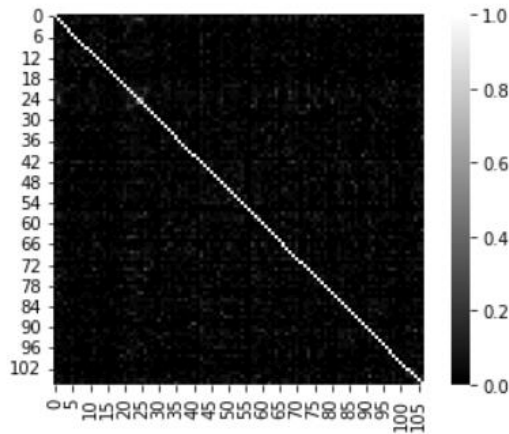


FIGURE 1. Cosine similarity heatmap of mission statements (microlevel).

The interpretation of the cosine similarity matrix of mission statements is conducted at both the microlevel (individual companies) and macrolevel (industries). At the microlevel, the heatmap (Figure 1) has an obvious diagonal line, as mission statements are naturally identical to themselves (i.e., cosine similarity values are equal to 1). In general, however, there is very

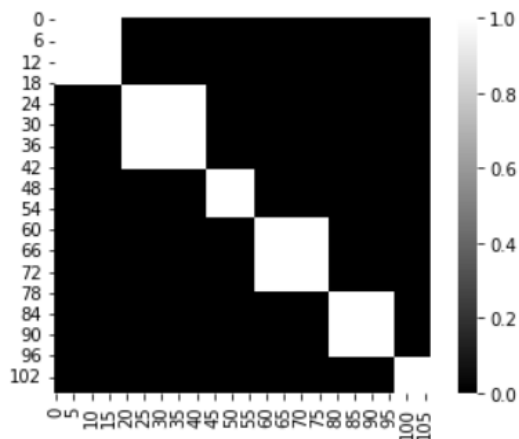


FIGURE 2. Cosine similarity heatmap of mission statements (extreme example).

little similarity between individual companies' mission statements (i.e., near-zero cosine similarity values).

As for the macrolevel, Figure 1 does not demonstrate any apparent similarities between the mission statements of companies operating in the same industry. To illustrate this, Figure 2 provides an extreme example of what the heatmap would look like if all companies' mission statements were primarily similar only to those within their respective industries.

To further investigate potential collective similarities in mission statements, we categorized all the cosine similarity values of mission statements into groups based on their respective industries. Next, we manually selected sets of values, considering the two industries being compared, and computed the mean value for each set. To enhance clarity, we rounded these average values and multiplied them by 1,000. Finally, we constructed a new heatmap (Figure 3) consisting of these averaged cosine similarity values, organized by industry.

Therefore, Figure 3 illustrates the degree of similarity among six industries in the Slovenian market, where the highest similarity values are once again concentrated along the diagonal line, suggesting that industries tend to be more similar to themselves. Nevertheless, we acknowledge that the statistical significance of this correlation can be considered, at best, only moderate.

The results of CDA reveal that Slovenian companies commonly employ mission statements characterized by vocabulary and grammar rich in experiential and expressive values. These mission statements mainly draw upon the authors' personal knowledge, beliefs, and social identity. This is reflected by the classic classification schemes primarily related to the locations (e.g., "... milk from East Slovenia region ...", "... Slovenian motorway network ...") and qualities (e.g., "... premium quality and healthy dairy products ...", "... high-quality electronic locking and access control systems ...", "... ensure safety and comfort ...") of products and services.

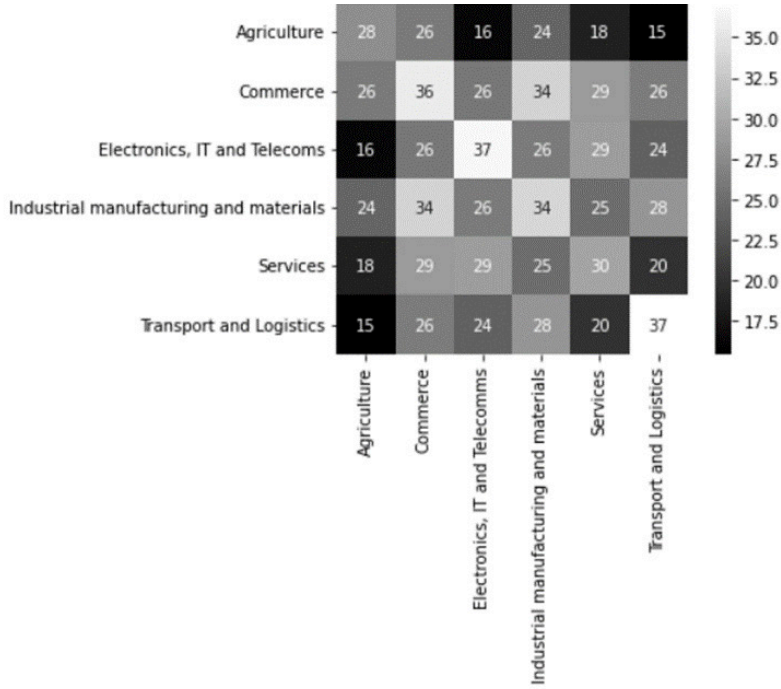


FIGURE 3. Cosine similarity heatmap of mission statements (macrolevel).

Furthermore, they provide clear definitions of the key processes and agency (the company itself). Additionally, these mission statements tend to employ active and positive sentence structures.

However, these mission statements frequently neglect relational values, making them less relatable to their intended audience. Stakeholders other than external parties are rarely mentioned explicitly, with even customers remaining vaguely defined. As a result, mission statements appear to be heavily centred on the company itself, as evidenced by the absence of pronouns other than “we” (referring to the company) and lack of personalization.

Moreover, the interpretation of a company’s purpose can vary among participants based on the situational and intertextual context. These mission statements often draw upon related discourse types, such as those on the company’s expertise and corporate values, leading to different customer understandings depending on their background knowledge and focus (e.g., product qualities or company values).

Lastly, the analysed mission statements are influenced by power relations at various levels, including organizational structure, market competition, and consumer attitudes and demands. While they address common social issues such as environmental responsibility (e.g., “... treasures of pure nature ...”, “... acting responsibly towards the environment ...”), traditions (e.g., “... rich tradition of milk production ...”), and modernization (e.g., “...

reliable and modern tailor-made solution ...”, “... taking modern approaches ...”), these concerns are often not emphasized, and existing power relations are not challenged, resulting in a relatively normative discourse.

## 4 Discussion

We acknowledge that the results of DA may lack consistency due to factors like analyst subjectivity, the dynamic nature of discourse, and a lack of standardized approaches. It must also be noted although mission statements may show correlations with business performance, they may not necessarily be the direct cause of it, and as such the significance of their presence or contents might be overstated.

We also recognize a significant language challenge with regard to SA and DA, as these analyses were conducted in English, whereas the mission statements were from Slovenian companies. As a result, we faced two options: either translating all the mission statements into English after collecting them, or solely collecting those mission statements that were originally written or translated into English by the business owners. We chose the latter option to avoid introducing any translation bias from our side. However, this approach does not completely eliminate the issue because English is not the primary language for Slovenian entrepreneurs. Thus, there remains a risk of losing or misinterpreting the original meaning in Slovenian and the potential for flawed translations.

Another major complication of the study lay in the difficulty of identifying and isolating the individual mission statements from the overall company description, particularly when they were not explicitly labelled as such. Nevertheless, we encountered many cases where one or more sentences closely resembled a mission statement in terms of conveying the company's purpose and primary objectives. We thus chose to consider these as valid mission statements to expand the sample size.

As a result, we ended up with a convenience sample, which comes with inherent disadvantages, including the possibility for selection bias. The subjective nature of selecting companies in this manner could potentially impact the generalizability and representativeness of our findings when extrapolated to the population of Slovenian companies as a whole.

The ability to extrapolate our findings beyond the scope of the study is further constrained by the relatively small sample size, which resulted in a limited number of companies and industries available for comparison.

In future research within this field, there are several key areas that could be explored to deepen our comprehension of mission statements and their significance for Slovenian companies. These areas encompass: investigating the impact of mission statements on employee motivation and effective leadership; conducting cross-cultural analyses of mission statements, expanding the scope beyond Slovenia to include other countries and cultures, especially within the European Union; and employing various combinations and types of DA and SA on a larger and more diverse sample of mission statements.

## 5 Conclusion

Slovenian companies predominantly use positive and active mission statements to effectively convey their purpose. However, these statements tend to be very conservative and often conform to the status quo. Thus, while some mission statements acknowledge modernization, sustainability, and other contemporary ideologies as positive attributes, they refrain from advocating for substantial change or an industry revolution, as such actions could disrupt existing power relations and introduce uncertainty.

To illustrate this, consider the original: “[...] to continue with its 75 years old, rich tradition of milk production [...] satisfying even the most demanding consumers home and abroad” and its revised version: “[...] to advance its 75 years of rich tradition in milk production [...] Apart from satisfying the most demanding consumers, we strive to support our local agricultural sectors by foresting a sustainable market. By implementing clean, environmentally friendly means of production, we challenge the current norms and contribute to the well-being of our planet and communities.”

In the same sense, compare: “By taking modern approaches and acting responsibly towards the environment [...]” with “In our commitment to redefine the future of transportation, we adopt innovative approaches and act responsibly towards the environment [...]”

Additionally, these statements mostly centre around the companies themselves, with limited mention of stakeholders and their requirements. As a result, the primary emphasis remains on the companies’ goals, values, and the qualities of their products and services.

For instance, compare the original: “We develop and implement [...] With the experience gained in many projects, we offer customers a reliable and modern tailor-made solution [...]” with the revised version: “We leverage our collective expertise in developing and implementing [...] The insights gained from numerous projects are not only directed towards reliable and modern solutions tailored to the customers’ needs but also provide an opportunity for the professional growth of our team members.”

Similarly, consider “[...] we optimize traffic flow and ensure safety and comfort on the Slovenian motorway network” and “It is within the expertise of our team of professionals to enhance the Slovenian motorway network with strategies that reflect our values, abide by regulations, and resonate with the community.”

Regarding the evaluation of similarity, the results were somewhat ambiguous. The analysis revealed that even though the mission statements may appear substantially distinct across individual companies, a moderate level of collective similarity emerges when they are categorized by industry.

Several potential factors might contribute to this phenomenon. On a smaller scale, variations in length, content, and vocabulary among mission statements could accentuate their diversity. In contrast, companies within the same industry often share similar objectives, products, target audiences, industry-specific norms, and standards. As a result, mission statements from organizations operating within the same sector tend to converge and exhibit resemblances, resulting in a certain level of collective similarity.

Overall, it is noteworthy that Slovenian Micro, Small and Medium Enterprises (MSMEs) and large organizations use mission statements mainly to articulate their objectives, products, and services. Furthermore, their intended message is primarily oriented towards external audiences rather than internal ones. It is as if the mission statement is perceived more as a marketing instrument than a managerial tool. As a result, by prioritizing external stakeholders, especially customers, over internal ones, these companies might be overlooking the potential advantages of mission statements. As argued by the studies presented earlier, these benefits are related to employee motivation and the development of effective leadership.

## References

- Alegre, Inés, Jasmina Berbegal-Mirabent, Adrián Guerrero, and Marta Mas-Machuca. 2018. "The real mission of the mission statement: A systematic review of the literature." *Journal of Management & Organization* 24 (4): 456–73. <https://doi.org/10.1017/jmo.2017.82>.
- Babnik, Katarina, Kristijan Breznik, Valerij Dermol, and Nada Trunk Širca. 2014. "The mission statement: Organisational culture perspective." *Industrial Management & Data Systems* 114 (4): 612–27. <https://doi.org/10.1108/IMDS-10-2013-0455>.
- Bart, Christopher K., and Mark C. Baetz. 1998. "The relationship between mission statements and firm performance: An exploratory study." *Journal of Management Studies* 35 (6): 823–53. <https://doi.org/10.1111/1467-6486.00121>.
- Bart, Christopher K., and Maureen Hupfer. 2004. "Mission statements in Canadian hospitals." *Journal of Health Organization and Management* 18 (2): 92–110. <https://doi.org/10.1108/1477260410538889>.
- Breznik, Kristijan. 2012. "Ranking and visualizing the keywords in mission statements of Slovenian companies." In *Management, Knowledge and Learning International Conference 2012 (MakeLearn)*, 969–75. Celje: International School for Social and Business Studies.
- Choque Aliaga, and Osman Daniel. 2018. "'The discontinuity in the continuity.' Michel Foucault and the archaeological period." *Topologik: Rivista Internazionale Di Scienze Filosofiche, Pedagogiche e Sociali* 23 (July): 1–15.
- David, Fred R. 2011. *Strategic Management: Concepts and Cases*. Upper Saddle River, NJ: Prentice Hall.
- Dermol, Valerij. 2013. "Relationship between mission statement and company performance." *Annals of the Alexandru Ioan Cuza University – Economics* 59 (1): 321–36. <https://doi.org/10.2478/v10316-012-0022-9>.
- Desmidt, Sebastian, Anita Prinzie, and Adelen Decramer. 2011. "Looking for the value of mission statements: A meta-analysis of 20 years of research." *Management Decision* 49 (3): 468–83. <https://doi.org/10.1108/00251741111120806>.
- Dijk, Teun A. van. 2014. *Discourse and Knowledge: A Sociocognitive Approach*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9781107775404>.
- Fairclough, Norman. 2013a. *Critical Discourse Analysis: The Critical Study of Language*. 2nd ed. New York: Routledge. <https://doi.org/10.4324/9781315838250>.
- . 2013b. *Language and Power*. 2nd ed. New York: Routledge.
- Gee, James Paul. 2011. *An Introduction to Discourse Analysis: Theory and Method*. 3rd ed. New York: Routledge. <https://doi.org/10.4324/9780203847886>.
- Gorenak, Mitja, and Suzana Košir. 2012. "The importance of organizational values for organization." In *Management, Knowledge and Learning International Conference 2012 (MakeLearn)*, 563–69. Celje: International School for Social and Business Studies.
- Hodges, Brian David, Ayelet Kuper, and Scott Reeves. 2008. "Discourse Analysis." *BMJ* 337: 570–75. <https://doi.org/10.1136/bmj.a879>.
- Johnstone, Barbara. 2008. *Discourse Analysis*. 2nd ed. Malden: Blackwell. <https://doi.org/10.1017/S0047404509090460>.

- Klemm, Mary, Stuart Sanderson, and George Luffman. 1991. "Mission statements: Selling corporate values to employees." *Long Range Planning* 24 (3): 73–78. [https://doi.org/10.1016/0024-6301\(91\)90187-S](https://doi.org/10.1016/0024-6301(91)90187-S).
- Manning, Christopher D., Prabhakar Raghavan, and Hinrich Schütze. 2009. *An Introduction to information retrieval*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511809071>.
- McHoul, Alec, and Wendy Grace. 2015. *A Foucault Primer: Discourse, Power and the Subject*. London: Routledge. <https://doi.org/10.4324/9780203501122>.
- McKinney, Wes. 2010. "Data structures for statistical computing in Python." *Proceedings of the 9th Python in Science Conference* 445: 56–61. <https://doi.org/10.25080/Majora-92bf1922-00a>.
- Pearce II, John A., and Fred David. 1987. "Corporate mission statements: The bottom line." *Academy of Management Executive* 1 (2): 109–15. <https://doi.org/10.5465/AME.1987.4275821>.
- Pedersen, Ove K. 2009. *Discourse Analysis. Working Paper no 65*. Frederiksberg: International Center for Business and Politics, Copenhagen Business School.
- Pedregosa, Fabian, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, Jake Vanderplas, Alexandre Passos, David Cournapeau, Matthieu Brucher, Matthieu Perrot, and Édouard Duchesnay. 2011. "Scikit-Learn: Machine learning in Python." *Journal of Machine Learning Research* 12 (85): 2825–30.
- Powers, Edward L. 2012. "Organizational mission statement guidelines revisited." *International Journal of Management & Information Systems (IJMIS)* 16 (4): 281–90. <https://doi.org/10.19030/ijmis.v16i4.7304>.
- Python. 2022. *Python Release Python 3.10.5*. <https://www.python.org/downloads/release/python-3105/>.
- Qaiser, Shahzad, and Ramsha Ali. 2018. "Text mining: Use of TF-IDF to examine the relevance of words to documents." *International Journal of Computer Applications* 181 (1): 25–29. <https://doi.org/10.5120/ijca2018917395>.
- Rahutomo, Faisal, Teruaki Kitasuka, and Masayoshi Aritsugi. 2012. "Semantic Cosine Similarity." *The 7th International Student Conference on Advanced Science and Technology* 4.
- Salem Khalifa, Azaddin. 2011. "Three Fs for the mission statement: What's next?" *Journal of Strategy and Management* 4 (1): 25–43. <https://doi.org/10.1108/17554251111110104>.
- Sheng, Yan-xiu, Zhi-qiang Wei, and Yong-Quan Yang. 2018. *Research of Text Classification Based on Improved TF-IDF Algorithm*. <https://doi.org/10.1109/IRCE.2018.8492945>.
- SloExport. 2022. *Database of Slovenian Exporters*. <https://www.sloexport.si/en/>.
- Stallworth Williams, Linda. 2008. "The mission statement: A corporate reporting tool with a past, present, and future." *The Journal of Business Communication* 45 (2): 94–119. <https://doi.org/10.1177/0021943607313989>.
- Taylor, Stephanie. 2013. *What Is Discourse Analysis?* London: Bloomsbury Academic.
- Waskom, Michael L. 2021. "Seaborn: Statistical data visualization." *Journal of Open Source Software* 6 (60): 3021. <https://doi.org/10.21105/joss.03021>.



## Attitudes of the Hungarian-American Diaspora to the Officialization of English in the United States

### ABSTRACT

The paper examines the attitudes and opinions of the representatives of Hungarian-American communities in the United States regarding the officialization of English at the federal level. The corpora of the analysis have been built with the help of a) the websites of Hungarian-American organizations in the United States; b) the legislative database of the US Federal Congress; and the digitized versions of the printed newspapers and magazines published in Hungarian in the United States. The method of classifying the attitude patterns relies on Terrence G. Wiley's language policy classification spectrum (1999) and on Richard Ruíz's classic "orientations in language planning" framework (1984), expanded by Hult and Hornberger (2016). The results indicate that while Hungarian-Americans have mostly been trying to maintain their first language (in the private domain), they have also embraced English and willingly assimilated into American society, resenting the alleged ethnolinguistic separatism of Hispanic Americans and their demands for special minority-language accommodations.

**Keywords:** Hungarian diaspora, United States, language ideologies, Official English

## Odnos madžarsko-ameriške diaspore do angleščine kot uradnega jezika v Združenih državah Amerike

### IZVLEČEK

Članek obravnava stališča in mnenja predstavnikov madžarsko-ameriških skupnosti v Združenih državah Amerike do angleščine kot uradnega jezika na zvezni ravni. Korpusi za analizo so bili oblikovani s pomočjo a) spletnih strani madžarsko-ameriških organizacij v Združenih državah Amerike; b) zakonodajne baze Zveznega kongresa ZDA; in digitaliziranih različic tiskanih časopisov in revij, ki izhajajo v madžarskem jeziku v Združenih državah Amerike. Metoda razvrščanja vzorcev stališč se opira na klasifikacijski spekter jezikovne politike Terrencea G. Wileyja (1999) in na klasični okvir "usmeritev pri jezikovnem načrtovanju" Richarda Ruíza (1984), ki sta ga razširila Hult in Hornberger (2016). Rezultati kažejo, da so Američani madžarskega porekla sicer večinoma poskušali ohraniti svoj prvi jezik (na zasebnem področju), vendar so sprejeli tudi angleščino in se voljno asimilirali v ameriško družbo, pri čemer so zavrnili domnevni etnolingvistični separatizem Hispanoameričanov in njihove zahteve po posebnih jezikovnih prilagoditvah za manjšine.

**Ključne besede:** madžarska diaspora, Združene države Amerike, jezikovne ideologije, uradna angleščina

# 1 Introduction

The size of the Hungarian-speaking diaspora in the United States has been dwindling fast for decades, as indicated by the relevant decennial censuses and the more recent American Community Survey statistics. Rapid assimilation into the mainstream is also evidenced by the concomitant Hungarian ancestry figures in census reports. These tendencies give rise to the hypothesis that embracing the English language has largely been taking (tacit) precedence over conscious and coordinated heritage language maintenance efforts in these communities, which may even have resulted in a general distrust of (additive) bilingual education programmes and the notion of language rights as well. Consequently, out of the three core characteristic features of a “diaspora” (as identified by Brubaker 2005, 5): dispersion in space; orientation to a ‘homeland’; and boundary-maintenance, at least the latter appears to be eroding quickly in view of the rapid language shift.

Relying on Bernard Spolsky’s tripartite language policy definition, comprising the language practices of a community, the (shifting) beliefs or ideological expectations (i.e., opinions about “what should be done”), and the language management efforts that frequently try to approximate practices to language-related beliefs (2004, 14; 2009, 4; 2021, 9), this analysis attempts to map the (English) language-related attitudes and potential attitudinal shifts in the Hungarian-American diaspora. This goal is to be reached by analysing the explicitly expressed opinions of the representatives or members of these communities concerning the officialization of English in the US, which is often regarded to be the most basic (and potentially most divisive) aspect of language management (Pułaczewska 2015, 3). I accept Graedler’s argument (based on Blommaert and Verschueren 1998) that newspapers “contribute to making issues of language and language use visible in the public sphere and are presumed to play a substantial role in the expression and mediation of a society’s language attitudes” (2014, 296). In order to compensate for potential biases, this exploration also focuses on the homepages of numerous Hungarian-American associations and on the legislative database of the US Federal Congress to find further Official English-related remarks by Hungarian-American contributors. “Attitudes” are understood here very broadly as “people’s positive or negative opinions or feelings about something” (Graedler 2014, 295), and are supposed to be rooted in deeper “orientations” or dispositions toward languages and linguistic diversity (discussed in detail in Section 4 below).

## 2 Hungarian-American Communities in the United States: Immigration Waves and Linguistic Assimilation

One of the most comprehensive sociolinguistics overviews of the Hungarian language in the United States is provided by Fenyvesi (2005, 265–318), who also reviews the successive waves of Hungarian immigration, which began in earnest in the 1880s (mostly due to economic hardship) and continued largely uninterrupted until World War I. During these decades, Hungarian-speaking immigrants – who typically found employment in mines and heavy industry – settled in the northeastern United States, primarily in Ohio, Pennsylvania, New York, Illinois, Michigan, Wisconsin and Indiana (Fenyvesi 2005, 267). Later, the immigration “Quota Acts” of the 1920s drastically restricted the influx of newcomers from Central, Eastern

and Southern Europe to the US, putting an end to the predominantly “peasant exodus” (Gazsó 2020, 52). The post-World War I boundary adjustments – resulting in severe territorial losses for Hungary – pushed many of the often forcibly relocated ethnic Hungarians from the successor states overseas, mostly to Australia, Canada and Latin-America (Gazsó 2020, 59).

After World War II, tens of thousands of “displaced persons”, refugees from the Communist regime, and “56-ers” after the failed revolution of 1956, found their new home in the US. (Before the 1920s, many Hungarians were of the “sojourner” type, who worked in the US for a few years, saved enough money to start a new life in Hungary, and eventually returned to their country of origin.) Borbándi’s periodization of the post-World War II Hungarian emigration waves identifies several milestones up to the mid-1980s that represented the peak years of the (partially US-bound) exodus (2006, 6). The Soviet occupation in 1945, the consolidation of Communist power in 1947 and the crushed revolution of 1956 forced hundreds of thousands of people to leave the country; however, the gradual softening of the dictatorial system after 1963 and the peak years of détente around 1975 – which also witnessed the return of the Holy Crown from the United States in 1978 – contributed to the cautious normalization of relations between the US and Hungary (Borbándi 2006, 6).

However, travel was extremely restricted (and immigration was practically prohibited) during the Communist years. After the fall of the Iron Curtain, new opportunities appeared for both temporary workers or would-be immigrants; however, today the primary destination for Hungarians in this regard is not the US but western and northern European countries, mostly within the EU.

The actual size of the Hungarian-speaking diaspora in the US is difficult to determine due to several factors. The 1890 census asked about the languages spoken in the US for the first time (United States Census Bureau 2021a), but this policy and the related questions have only been consistent (and the results comparable in a more-or-less reliable manner) since 1980 (United States Census Bureau 2021b). Even now, the self-reports based on the three relevant questions (“Does this person speak a language other than English at home?”; “What is this language?”; “How well does this person speak English (very well, well, not well, not at all)?”) reveal very little about the frequency of minority language use (in the private domain). Furthermore, we can only guess at the actual levels of heritage language proficiency, as the survey focuses only on the English language in this regard.

However, since 2005 the English proficiency and home language use-related questions have been transferred from the decennial census to the American Community Survey (ACS) (Zhou 2015, 209). The ACS is an ongoing survey that provides information on a yearly basis (United States Census Bureau 2022), but it is sent only to a sample of about 3.5 million addresses in the country – as opposed to the once-a-decade census, which attempts to count and gather information on every person living in the United States at the same time (United States Census Bureau 2023).

Consequently, the most recent decades are relatively well-documented as far as minority language use is concerned. Unfortunately, the survey figures indicate a steadily shrinking Hungarian-speaking diaspora in the US: while in 2010 1,501,736 people reported Hungarian

ancestry (United States Census Bureau 2010b), only 94,464 answered that they actually used Hungarian (at least to a certain degree) at home (United States Census Bureau 2010a). By 2021, the corresponding figures were 1,221,273 for the national origin question (United States Census Bureau 2021d), representing a 19% decrease, and 68,716 for home language use (United States Census Bureau 2021c) indicating a 23% drop within a decade.

### 3 The Officialization of English in the United States

Although the recorded history of *de jure* officialization efforts at state level dates back to the immediate post-World War I period, the more recent (and increasingly successful) Official English legislative attempts have been characterizing and sometimes dominating the language policy scene since the 1980s. In 2023, *ProEnglish* (a leading advocate of the issue) listed 32 states with symbolic or (in recent years, in particular) substantive and enforceable officialization policies (*Official English Map*).

At the federal level, the formal beginnings of the “Official English” movement are associated with Senator S.I. Hayakawa’s ambitious amendment proposal to the US Constitution in the form of a joint resolution (S.J.Res. 72, 97th Cong., 1st Sess.), introduced on April 27, 1981. Although the Hayakawa resolution failed, it catapulted the issue into the national limelight, and to date there have been more than one hundred similar bills and resolutions introduced in Congress – although none of them have been passed.

The proposal that came closest to being enacted was the Emerson “English Language Empowerment Act” (H.R. 123 EH, 104th Cong., 2nd Sess.), which reached the Congressional floor on August 1, 1996, and was passed by the House of Representatives. However, it was not even introduced in the Senate as then-President Bill Clinton promised to veto the measure, should it have been passed by the upper chamber as well.

The current flagship Official English proposal in the House is the “English Language Unity Act of 2023” (H.R. 997 IH, 118th Cong., 1st Sess.), and its identical sister bill in the Senate (S. 1109 IS, 118th Cong., 1st Sess.). If passed – the chances of which are negligible – the “English Language Unity Act of 2023” would declare English the official language of the United States, and “establish a uniform English language rule for naturalization”, including the requirement that “All citizens should be able to read and understand generally the English language text of the Declaration of Independence, the Constitution, and the laws of the United States” (H.R. 997, 118th Cong., 1st Sess., Sec. 3, “Ch. 6, “§164, “(a)).

According to James Crawford, since the 1990s “Official English” has become a “mainstream phenomenon” (2000, 4–5), a symbol of patriotism. This observation resonates with Deborah Schildkraut’s argument that a fuller elaboration and explanation of American identity requires the extension of the traditional but limiting interpretation of the essence of the concept, which is often simplistically equated it with “liberalism” – describing the United States as the “land of freedom and opportunity” (2005, 6). At the very minimum, Schildkraut proposes three other, frequently neglected aspects (“traditions”) of American identity to be taken into consideration, challenging the more or less traditional, liberal interpretation. These are the civic republican tradition (emphasizing the duties and responsibilities of the citizens, not only

the potential opportunities and benefits); the ethnocultural, “white Anglo-Saxon Protestant” tradition; and the incorporationist tradition (offering a middle ground between melting pot-style assimilation and minority cultural maintenance) (Schildkraut 2005, 6).

As far as Official English policies are concerned, the proponents of the liberal tradition are most likely to oppose official-language laws (e.g., for potentially violating the freedom of speech constitutional guarantee), while civic republicanism emphasizes the need to sustain a sense of community, and thus is more likely to endorse attempts to promote the national language (Schildkraut 2003, 473–74). Restrictions on minority language use are not alien to the ethnocultural tradition, either (Schildkraut 2003, 474); whereas incorporationism is more ambiguous in this respect, as it “celebrates ethnic diversity and praises maintaining cultural traditions while also supporting assimilation and the emergence of new, uniquely American traditions” (Schildkraut 2003, 493).

#### 4 Aims, Method and Corpora

The current analysis focuses on the documented attitudes (and possible attitudinal shifts) of Hungarian-Americans toward the officialization of English, and tries to determine whether these manifestations represent more-or-less consistent language ideologies – i.e., “commonsense notions about the nature of language” (Woolard 1992, 235) and about the “proper” role of languages and language varieties in a community.

Several scholars have pointed out that popular, language-related beliefs in the United States are often dominated by strong, assimilation-oriented expectations, despite the routinely diversity-praising rhetoric (Macías 2000, 53; Wiley 2000, 84; Baron 2004; Cutshall 2004/2005). An earlier review of the relevant literature identified three major sets of arguments that dominate language ideology-related debates in the United States (Czeglédi 2008, 35–57):

- (1) The (mis)interpretation of (immigrant) history in general and that of salient events or personal histories in particular;
- (2) Unfounded beliefs about (second) language acquisition and the effectiveness of bilingual education;
- (3) The extent to which the English (and exclusively the English) language is considered to be a/the key component of American identity and nationhood.

Obviously, it is not a coincidence that these categories overlap considerably with the major areas of language policy conflicts in the United States as identified by Schmidt (2000, 11): 1. “Bilingual” education (which may often mean the implementation of almost wholly English-monolingual models); 2. Ensuring linguistic access to (government) services (which is frequently equated with language rights guarantees); 3. Declaring English the (sole) official language of the United States (either in the form of a substantive constitutional amendment or bill; or by a symbolic concurrent resolution).

Attitudes or “orientations” towards policy proposals in these areas (and towards linguistic diversity, languages and language communities in general) can be classified in several ways, two of which have been used in the present analysis.

In 1984, Richard Ruíz outlined three possible orientations – language-as-problem, language-as-right, and language-as-resource – toward languages in society (1984, 15–34), defined as complex, unconscious and mostly pre-rational dispositions, “related to language attitudes in that they constitute the framework in which attitudes are formed” (1984, 16). Francis M. Hult and Nancy H. Hornberger (2016) elaborated further on Ruíz’s orientations by fine-tuning and expanding the original characteristics of each perspective. Accordingly, the language-as-problem orientation typically values monolingualism in the dominant language and sees minority languages as an overall threat – consequently, the quick mainstreaming of minority students into the dominant language is a desirable goal (Hult and Hornberger 2016, 33). Language-as-right focuses on granting linguistic access to government services, civil rights and to society in general, ideally involving legal guarantees against discrimination (negative rights), as well as codified rights allowing or even encouraging the use of the minority languages in specific domains (positive rights) (Hult and Hornberger 2016, 33). The language-as-resource orientation values societal multilingualism and cultural diversity. Here, languages are thought to have both extrinsic (societal) and intrinsic (individual) value (Hult and Hornberger 2016, 33). Consequently, minority language maintenance efforts and foreign language learning/teaching are both encouraged.

Another easily applicable yet highly informative framework for formal language policy analysis has been developed by Terrence G. Wiley since the late 1990s (Wiley 1999, 21–22; Wiley and de Korne 2014, 1–2). Wiley classifies the full range of possible policies according to a spectrum of categories containing a promotion-, expediency-, tolerance-, restriction- or repression-orientation. Official English proposals are naturally considered to be promotive from the perspective of the English language; however, they are frequently restrictive (or could even be repressive) for minority languages (Czeglédi 2008, 101–4; 111–15). Expediency denotes short-term minority language accommodations – e.g., the provision of court interpreters, bilingual ballots, transitional bilingual education – which are not intended to foster minority-language maintenance (Johnson 2013, 35).

Both Ruíz’s and Wiley’s classification schemes have been used to determine the language-related attitudes of Hungarian-American minorities toward the English language, the Hungarian heritage language, and toward other minority languages in the United States.

Three corpora have been analysed for this purpose: 1. The homepages of Hungarian-American associations and clubs in the United States; 2. The legislative database of the Federal Congress; 3. The newspapers and magazines issued in the US by the Hungarian diaspora (in Hungarian). In all three sources, I have focused on those documents and articles that contained either the keywords “official English” AND “Hungar\*” (in the English-language documents) or “hivatalos nyelv” [“official language”] AND “angol” [“English”] (in the Hungarian-language sources).

As the newspaper corpus turned out to be the most informative, keyword searches focusing on the other two main language policy battlegrounds – bilingual education [“kétnyelvű oktatás”] and language rights [“nyelvi jogok”] – were also carried out to corroborate (or disprove) the findings related to the issue of officialization.

## 5 Findings and Discussion

### 5.1 The Homepages of Hungarian-American Associations and Clubs in the United States

The largest Hungarian-American umbrella organization in the United States is the American Hungarian Federation (AHF) (“Amerikai Magyar Szövetség,” AMSZ in Hungarian), originally founded in 1906 in Cleveland, Ohio (The American Hungarian Federation 2023a). The homepage is available in the English language only, which indicates high levels of language shift among the potential target audience. The Federation makes no secret of the fact that “today there are many Hungarian-Americans who feel strongly about their heritage but may not speak the language” (The American Hungarian Federation 2023a). Inclusivity and reaching out to second, third-, etc. generation Hungarian-Americans thus require an approach that does not alienate those who have never had the chance to learn Hungarian. While the American Hungarian Federation has been supporting “educational and cultural preservation programs” through the Hungarian American Education and Cultural Fund (“Amerikai Magyar Oktatási és Kulturális Alap” (AMOKA)), these are mostly focused on helping the language and cultural maintenance of ethnic Hungarians living in the neighbouring countries, separated from the mother country by the post-World War I boundaries (The American Hungarian Federation 2023c).

The AHF homepage lists 70 member organizations and Hungarian-American clubs throughout the United States, 23 of which have linked homepages. Out of these 23, only four have at least some of their online content available in Hungarian. These are the Cleveland-based “Hungarian Association”/Magyar Társaság (<https://www.hungarianassociation.com/>); the Hungarian Club of Colorado (<https://huclub.org/>); the Hungarian Society of Massachusetts (<https://bostonhungarians.org/>); and the Metroplex Magyar Cultural Circle (in Texas) (<https://magyarszo.net/>).

The AHF has no resources related to the officialization of English in the US at all. The “official language” search string returned five, slightly outdated records from their database: four of which discuss the “Draconian language law” in Slovakia, which restricted the use of Hungarian in official domains (see, e.g., The American Hungarian Federation 2023b).

While the AHF clearly (and understandably) regards the Hungarian language as an invaluable asset in its minoritized, autochthonous context in the successor states to the former Austro-Hungary, the Federation appears to be somewhat more reluctant to mobilize its (definitely not unlimited) resources to delay (the presumably inevitable) language shift in the American melting pot.

### 5.2 Officialization-Related Attitudes in the Legislative Database of the Federal Congress

The online congress.gov database search returned five relevant records for the “official English” AND “Hungar\*” query, mostly from the 1990s. In all of these contexts, assimilation-oriented opinions were expressed either directly or by referring to famous Hungarian-Americans who fulfilled the American Dream with the help of the English language.

The first such instance can be traced back to the Congressional Record in 1991, when the contemporary flagship Official English bill, the “Language of Government Act” (introduced by Representative Bill Emerson, R-MO) was discussed (or rather, promoted) by Rep. Emerson himself and his colleagues. Representative William L. Dickinson (R-AL) argued for officialization by enumerating several immigrant success stories which were made possible “only in America” (and only in English), e.g., the illustrious career of former Senator Samuel I. Hayakawa, the sponsor of the first recorded Official English proposal at the federal level in 1981, and that of Arnold Schwarzenegger, who was already one of the best-known movie stars in the early 1990s. Rep. Dickinson also recalled the story of a (supposedly) Hungarian immigrant, Stephen Baker, who became a highly successful advertising executive, and created some of the most memorable ad campaign slogans of the 1960s and 1970s, including “Let your fingers do the walking” for AT&T’s Yellow Pages (Dickinson 1991, 18876). In fact, Baker was born in Vienna as Steven Bacher, although he did indeed spend “much of his childhood” in Budapest (Bayot 2004).

The second context in which references to Hungarian-American minorities were recorded was the Elementary and Secondary Education Act (ESEA) reauthorization debate in 1994. Eventually, the Improving America’s Schools Act (IASA) that was passed in 1994 turned out to be the most language-as-resource oriented reauthorization of the ESEA (Czeplédi 2017, 15). However, it was weakened significantly in this regard after the “Republican Revolution of 1994”, when the Republican-dominated Congress quickly slashed funding for Title VII (the Bilingual Education Act, or BEA), and (abortive) efforts were made to rescind the BEA altogether (Schmidt 2000, 18).

Representing the assimilationist side, Congressman Toby Roth (R-WI) argued during the March 21 debate that immigrant children were retained too long in bilingual education programmes, which – according to him – paved the way toward the “disuniting of America” (referring to the title of Arthur Schlesinger’s critical reflections on multiculturalism, published a few years before), resulting in the US “turning into another Canada or another Yugoslavia” (Roth 1994, 5736). Rep. Roth also added that earlier he had a hearing on Capitol Hill with people of “Spanish background”, immigrants from the Soviet Union and Hungary, and “all of them opposed to bilingual education because it harms their people” (1994, 5736).

The most characteristic, firsthand Hungarian-American opinions related to officialization were recorded in 1995–96, during the Congressional hearings about the Bill Emerson “English Language Empowerment Act” (H.R. 123). The earlier version of the proposal (then called “The Language of Government Act of 1995”) was introduced in the House of Representatives on January 4, 1995 by Rep. Emerson, and was referred to the Committee on Economic and Educational Opportunities, then to the Subcommittee on Early Childhood Youth and Families. As there were several other Official English bills introduced in Congress in the very same year, all of them were discussed at two subcommittee hearings regarding “English as a Common Language”; the first of which occurred on October 18, 1995, and the second on November 1, 1995 (H.R. Rep. No. 104–723, 3–4 (1996)). During the second hearing, one of the witnesses was “Mr. Charles Gogolak, former professional football player and Hungarian immigrant” (H.R. Rep. No. 104–723, 4 (1996)). Although his testimony was not recorded,

Charles Gogolak has been an ardent supporter of Official English ever since – as evidenced by his several decades’ long membership on the Advisory Board of US English – together with, for example, Arnold Schwarzenegger and Francis Fukuyama (US English 2022). George Dozsa, the president of an old Hungarian-American organization, the Hungarian Reformed Federation of America (a fraternal life insurance society organized in Trenton, New Jersey, in 1896), also expressed their “full support of the goals and initiatives of the official English language policies with specific regard to H.R. 123, the Official Language of Government Act” in a letter sent to the subcommittee (H.R. Rep. No. 104–723, 18 (1996)).

Debates about the “English Language Empowerment Act of 1996” elicited further contributions, even from a legislator with Hungarian roots. On August 1, 1996, Representative Ernest Istook (R-OK) expressed his unequivocal support for H.R. 123, and explained the origin of his surname (coming from the Hungarian “Istók” – the name of his immigrant grandfather), illustrating the necessity of melting pot-type assimilation patterns through his own family history (Istook 1996, H9763). Rep. Istook mentioned that his father was bilingual (using Hungarian only at home), but he himself was not a “hyphenated American” – rather a “real” American who believed that the common tongue should be English (Istook 1996, H9763).

All of these legislative records indicated that Hungarian immigrant experiences, histories, and personal (and political) opinions unanimously supported the officialization of the English language in the United States. Only the English language was regarded to be a nation-building asset, and minority languages were implicitly categorized as problematic (although, obviously no one argued against heritage language use in the private domain).

### 5.3 Officialization-Related Attitudes in the Hungarian-Language Newspapers Published in the United States

From the three major databases for this analysis, it was the Arcanum Digitheca archive that provided the deepest insight into the focal issue, despite the fact that many of the Hungarian-language publications were discontinued years or even decades ago – a fact that in itself illustrates how rapid the language shift has actually been, more-or-less since the immigration restrictions of the 1920s.

As of February 2023, the Arcanum Digitheca database contained the digitized versions of 22 Hungarian-language newspapers and magazines (published in the United States at some point) which have at least one reference to Official English throughout the decades. The oldest of these is *Amerikai Magyar Népszava* (originally established in 1891, in New York), which merged with the Cleveland-based *Szabadság* in 1948, and continued to appear in print until 2019 (Wikipedia 2023, par. 1). The digitized issues in Arcanum cover the years between 1909 and 2018. In addition to *Amerikai Magyar Népszava*, New York was the headquarters of 12 other relevant Hungarian-language publications, while Pittsburgh had one, Cleveland three, Chicago three, and Los Angeles two.

The “hivatalos nyelv” [“official language”] AND “angol” [“English”] query produced 31 relevant records between 1958 and 2007, the geographical distribution of which was the following:

TABLE 1. The geographical distribution of officialization-related articles in the newspapers published by the Hungarian diaspora in the United States.

City/Region	No. of Hungarian papers	Relevant articles (1958–2007)
New York, NY	13	14
Pittsburgh, PA	1	2
Cleveland, OH	3	1
Chicago, IL	3	3
Los Angeles, CA	2	11

New York and Los Angeles were the two cities/regions where Official English received the most journalistic attention. Here, two papers appeared to be the main publications in this regard: *Amerikai Magyar Népszava – Szabadság* (New York) printed twelve relevant articles, while *Californiai Magyarság* (Los Angeles) published nine.

The number of articles appearing in a given year largely reflected the intensity of the Official English debate: 1) the immediate aftermath of the introduction of Sen. Hayakawa's S.J.Res. 72 in 1981 (see Section 2 above); 2) California's own political debates about Proposition 63, the ballot measure which amended the state constitution in 1986 to declare English the official language there; 3) the partial success of H.R. 123 in 1995–96 (see Section 3 above); and the heightened "English-only" sentiments in 2006, which again raised the level of sponsorship behind the leading Official English bill in the House of Representatives to almost 38% – second only to H.R. 123's 45.3% a few years before (Czeglédi 2018, 90).

TABLE 2. The timeline of officialization-related articles in the newspapers published by the Hungarian diaspora in the United States.

1958	<i>Szabad Magyarság</i> (available issues: 1956-1962): New York			
1974	<i>Új Világ</i> (available issues: 1972-1993): Los Angeles			
1978	<i>Californiai Magyarság</i> (available issues: 1939-2009): Los Angeles			
1981	<i>Californiai Magyarság</i> (available issues: 1939-2009): Los Angeles			
1982	<i>Amerikai Magyar Népszava - Szabadság</i> (available issues: 1909-2018): New York	<i>Californiai Magyarság</i> (available issues: 1939-2009): Los Angeles	<i>Új Világ</i> (available issues: 1972-1993): Los Angeles	<i>Californiai Magyarság</i> (available issues: 1939-2009): Los Angeles
1983	<i>Amerikai Magyar Népszava - Szabadság</i> (available issues: 1909-2018): New York	<i>Californiai Magyarság</i> (available issues: 1939-2009): Los Angeles	<i>Amerikai Magyar Népszava - Szabadság</i> (available issues: 1909-2018): New York	<i>Amerikai Magyar Népszava - Szabadság</i> (available issues: 1909-2018): New York
1984	<i>Amerikai Magyar Népszava - Szabadság</i> (available issues: 1909-2018): New York	<i>Chicago és Környéke</i> (1958-1990): Chicago		
1985	<i>Californiai Magyarság</i> (available issues: 1939-2009): Los Angeles	<i>Californiai Magyarság</i> (available issues: 1939-2009): Los Angeles		
1986	<i>Szabadság</i> (available issues: 1971-1994): Cleveland	<i>Chicago és Környéke</i> (1958-1990): Chicago	<i>Californiai Magyarság</i> (available issues: 1939-2009): Los Angeles	
1987	<i>Chicago és Környéke</i> (1958-1990): Chicago			
1988	<i>Californiai Magyarság</i> (available issues: 1939-2009): Los Angeles			
1992	<i>Magyarság</i> (available issues: 1952-1997): Pittsburgh			
1994	<i>Magyarság</i> (available issues: 1952-1997): Pittsburgh			
1995	<i>Amerikai Magyar Népszava - Szabadság</i> (available issues: 1909-2018): New York	<i>Amerikai Magyar Népszava - Szabadság</i> (available issues: 1909-2018): New York		
1996	<i>Amerikai Magyar Népszava - Szabadság</i> (available issues: 1909-2018): New York	<i>Amerikai Magyar Népszava - Szabadság</i> (available issues: 1909-2018): New York		
1997	<i>Transsylvania - Erdélyi Tájékoztató</i> (available issues: 1959-2006)			
2006	<i>Amerikai Magyar Népszava - Szabadság</i> (available issues: 1909-2018): New York	<i>Amerikai Magyar Népszava - Szabadság</i> (available issues: 1909-2018): New York		
2007	<i>Amerikai Magyar Népszava - Szabadság</i> (available issues: 1909-2018): New York			

As far as officialization-related attitudes were concerned, all of these articles were wholeheartedly supportive of the idea, regardless of the year or place of publication. The only quasi-exception was an interview in *Amerikai Magyar Népszava – Szabadság* with the bilingual education expert Ramon Santiago from Georgetown University, who argued that the melting pot concept was already an outdated interpretation of American identity and nation-building, and the preservation of ethnic languages and cultures should take precedence (“Négyszemközt: Bevándorlás és kétnyelvűség” 1983, 10). Nevertheless, the (anonymous) author/interviewer did not hide their pro-assimilation bias (and not-so-slight grudge against mostly Latino immigrants) by pointing out the fact in the opening sentences that when the Hungarian readers of the article arrived in the States, nobody rushed to translate the official documents for them (“Négyszemközt: Bevándorlás és kétnyelvűség” 1983, 10). Furthermore, they added, as the (de facto) official language of the US was English, it was everyone’s duty to learn it.

Clearly, the second part of the article on “immigration and bilingualism” (“Négyszemközt: Bevándorlás és kétnyelvűség 2” 1983, 10), which was published by *Amerikai Magyar Népszava – Szabadság* one week later, shared the opinion of the invited expert, ex-Senator S.I. Hayakawa on several issues, including the perceived need to discontinue the translation of official documents, eliminate bilingual education, and – obviously – to declare English the official language of the country.

In many ways, the Hayakawa interview represented the views of the other Hungarian-language articles in the corpus regarding not only officialization but also bilingual education and minority linguistic access to government services, civil rights or other (re)sources. Both bilingual education and the access guarantees were seen as harmful for nation-building purposes and unfair to earlier immigrants (who could not take advantage of these services prior to the 1960s and 70s).

As far as minority linguistic access was concerned, the articles were doubtful or even dismissive of these accommodations. A 1982 editorial in *Amerikai Magyar Népszava – Szabadság* found it moderately shocking to see a Spanish-English bilingual edition of the Sears Roebuck catalogue published (alongside the Spanish-language editions of the *Chicago Sun Times* and the *Arizona Republic*), pointing out that “we (i.e., Hungarian-Americans) were not pampered” (“minket nem kényeztettek”) in this respect (“Minket nem kényeztettek...” 1982, 2). Zoltán Mikó Kovács remarked sarcastically in *Californiai Magyarság* a year later that there was no official language – that is why ballot papers were printed in two languages (Mikó Kovács 1983, 3). In 1995, Éva Náday briefly analysed the nature of the contemporary “language war” in the US, and concluded that many of the access-related minority language rights were absolutely against common sense: most glaringly, the possibility of taking the citizenship test in non-English languages (Náday 1995, 2).

The Náday article also focused on bilingual education, which was seen as the most heated battlefield in the language war, contributing to the emergence of “multicultural chaos” in the US (1995, 2). One and a half decades before, the columnist Dr. Endre Nánay had characterized bilingual education and its outcomes in rather simplistic terms. The author argued that as a result of educating schoolchildren in their native tongues the English

language would be relegated to a timetabled subject, and it was never going to be mastered at the necessary level, which could prevent minority students from obtaining tertiary education degrees (Nánay 1981, 1). According to Nánay, the most problematic immigrants were the Latinos, who refused to learn the English language and practiced ethnic separatism *en masse* – as opposed to Hungarians, who – despite cultivating their mother tongue at home – knew their obligations toward the English language as dutiful citizens. Furthermore, Hungarians had made significant sacrifices for the American ideal of freedom (by rising up against the occupying Soviet forces in 1956) even before arriving in the United States (Nánay 1981, 1).

A quarter of a century later, complaints about the declining quality of educational standards – attributed to the spread of bilingual education demanded mainly by Latinos – were featured in a few articles in the Hungarian-language press. One contributor, Tamás Devecseri described the experience of his son, who was learning first- or second-grade materials in mathematics in the fifth grade in a New Jersey elementary school, but he had Spanish lessons almost every day (Devecseri 2007, 8).

The Spanish language always appeared in a problem-oriented context, with frequent alarmist predictions that “the new immigrants were transforming American society” beyond recognition (“Az új emigránsok átalakítják...” 1986, 3), which was expected to be countered and remedied by the promotion and officialization of English.

The other minority language predominantly (mis)characterized as a historical threat to American nation-building was German. Between 1983 and 1994, the Hungarian papers published in the US uncritically repeated the “Muhlenberg legend” (i.e., the apocryphal story about German having missed becoming the official language of the US by only one vote after winning independence from Britain) not fewer than six times in the Official English context alone. The authors of these articles were Péter Halász (1983, 9); Dr. Endre Nánay (1985, 1); György Mikes (1986, 12); an unnamed author (“Milyen fontos egy szavazat?” 1992, 3); and Dr. Rezső Dabas (1994, 1). The only cautiously optimistic thought experiment related to the Muhlenberg legend was an article by Ottó Habsburg written in 1984, which – based on the assumption of the emergence of a German-speaking United States after independence – retroactively envisioned the peaceful (and permanent) secession of the Confederate states; the Great War ending in a draw; and the avoidance of a second global conflagration altogether (1984, 3).

With the exception of the previously mentioned Ramon Santiago interview in *Amerikai Magyar Népszava – Szabadság* (“Négyszemközt: Bevándorlás és kétnyelvűség” 1983, 10), minority languages were never considered to be a societal resource in general. The Hungarian language was mentioned twice as an invaluable cultural and linguistic asset for the Hungarian community (Metzger 1974, 3; Nánay 1981, 1), but in both cases this opinion was tempered by the simultaneous expression of unwavering loyalty to the United States and to the only language that was seen to bind the ever more diverse country successfully together: English.

Although there were 27 articles in the Hungarian-American press corpus that mentioned “bilingual education” at least once, the majority of these pieces expressed their concern about the lack of bilingual education opportunities offered to Hungarian autochthonous minorities in the neighbouring states (in Slovakia, Romania and Serbia) – at least at the time

of their publication. Altogether 12 articles discussed bilingual education policies in the US (including the ones mentioned above in the context of Official English), and all of these were overwhelmingly dismissive of such policies, especially from 1981, after the introduction of S.I. Hayakawa's Official English resolution in Congress. They typically emphasized the (allegedly) divisive nature of bilingual education ("A kétnyelvűséget ellenzi az új kormány" 1981, 2); raised the alarm about the danger of producing "semilingual illiterates" (Márer 1981, 10), and even urged immigration restrictions to stem these harmful tendencies (Dr. Ternovszky 1983, 10).

The few articles that touched upon the issue of language rights (bilingual ballots, translation services) were also extremely critical of these accommodations, especially when they were seen as unduly benefiting Hispanic Americans. There was not a single article among the examined pieces that could have been classified as a "language-as-right" (let alone "language as-resource") oriented piece, which might have argued for the taxpayer-funded promotion (or even preservation) of immigrant minority languages, or would have supported expediency-oriented policies to ease the transition into the English-speaking mainstream (e.g., multilingual ballots, court interpreters or translation services).

## 6 Conclusion

This attempt to trace the potential attitudinal shifts of the Hungarian-American communities toward the officialization of English has proved to be a partial failure – since instead of shifts (and even slight changes) it was the monolithic uniformity of views throughout the examined decades that dominated the relevant language ideologies.

The census figures have been showing rapid assimilation patterns among Hungarian-Americans, who appear to have overwhelmingly supported the officialization of English, and regarded the attainment of English proficiency as a civic republican duty and an unmistakable symbol of loyalty toward their new home. While the desire to maintain (and pass on) the heritage language and traditions was occasionally expressed in minority newspapers, these efforts turned out to be rather ineffective in the American context.

The examined sources generally agree that "the good old immigrants of yesteryear" (Hungarians included) assimilated into American society without problems – although the frequent references to the myth that Germans attempted to have their language established as the only official language at the end of the 18th century somewhat contradicted these idealized views of the past (when Latino immigrants were few and far between). For Hungarian-Americans, Latinos and their supposed unwillingness to Americanize themselves quickly enough, plus their "unjustifiable" demands for expediency-oriented Spanish-language access to government services and for (the "strong" forms of) bilingual education represented the signs of ethnolinguistic separatism and even blatant disloyalty toward the United States. It is also clear that bilingual education in particular was discussed extremely superficially in the relevant sources of the corpus: psycholinguistic arguments were entirely missing both from the congressional documents and the newspapers or magazines (let alone the homepages of Hungarian-American organizations).

Without doubt, the findings indicate that the key component of American identity formation and nation-building for the Hungarian diaspora was the English language, and any, relatively recent bilingual accommodations were seen as unwarranted and unjust – especially from the perspective of those minorities that could not have benefited from similar measures in the past.

## References

- “A kétnyelvűséget ellenzi az új kormány.” 1981. *Californiai Magyarság*, May 1, 1981, 2. Retrieved from Arcanum DigiTheca. [https://adt.arcanum.com/hu/view/CaliforniaiMagyarsag\\_1981/?query=%22k%C3%A9tnyelv%C5%B1+oktat%C3%A1s%22&pg=169&layout=s](https://adt.arcanum.com/hu/view/CaliforniaiMagyarsag_1981/?query=%22k%C3%A9tnyelv%C5%B1+oktat%C3%A1s%22&pg=169&layout=s).
- Arcanum DigiTheca. 2023. <https://adt.arcanum.com/hu/>.
- “Az új emigránsok átalakítják Amerika társadalmát.” 1986. *Chicago és Környéke*, July 12, 1986, 3. Retrieved from Arcanum DigiTheca. [https://adt.arcanum.com/hu/view/ChicagoEsKornyeke\\_1986/?query=az%20%C3%BAj%20emigr%C3%A1nsok%20%C3%A1talak%C3%ADt%C3%A1k&pg=330&layout=s](https://adt.arcanum.com/hu/view/ChicagoEsKornyeke_1986/?query=az%20%C3%BAj%20emigr%C3%A1nsok%20%C3%A1talak%C3%ADt%C3%A1k&pg=330&layout=s).
- Baron, Dennis. 2004. “Spanish, English and the New Nativism.” <https://debaron.web.illinois.edu/essays/native.htm>.
- Bayot, Jennifer. 2004. “Stephen Baker, Ad Executive, Dies at 83.” *The New York Times*, September 24, 2004. <https://www.nytimes.com/2004/09/24/obituaries/stephen-baker-ad-executive-dies-at-83.html>.
- Blommaert, Jan, and Jef Verschueren. 1998. “The role of language in European nationalist ideologies.” In *Language Ideologies: Practice and Theory*, edited by Bambi B. Schieffelin, Kathryn A. Woolard and Paul V. Kroskrity, 189–201. Oxford: Oxford University Press.
- Borbándi, Gyula. 2006. *A magyar emigráció életrajza 1945–1985*. Hungarian Electronic Library. <https://mek.oszk.hu/03400/03472/html/emigr1.htm#>.
- Brubaker, Rogers. 2005. “The ‘diaspora’ diaspora.” *Ethnic and Racial Studies* 28 (1): 1–19. <https://doi.org/10.1080/0141987042000289997>.
- Crawford, James. 2000. *At War with Diversity: US Language Policy in an Age of Anxiety*. Clevedon: Multilingual Matters.
- Cutshall, Sandy. 2004/2005. “Why We Need ‘The Year of Languages.’” *Educational Leadership* 62 (4): 20–23. [http://www.ascd.org/publications/ed\\_lead/200412/cutshall.html](http://www.ascd.org/publications/ed_lead/200412/cutshall.html)
- Czeglédi, Sándor. 2008. *Language Policy, Language Politics, and Language Ideology in the United States*. Veszprém: Pannonian University Press.
- . 2017. “Promotion, tolerance or repression? The treatment of minority, immigrant and foreign languages in the elementary and secondary education acts.” *Hungarian Journal of Applied Linguistics (HJAL)* 12 (1): 1–18. <https://doi.org/10.18460/ANY.2017.1.002>.
- . 2018. “The U.S. Federal Congress and the officialization of English: Arguments, ideas and ideologies in federal-level legislative proposals. In *Az emberi sors és a történelem kereszteződésében: tanulmánykötet Frank Tibor 70. születésnapjára/At the Crossroads of Human Fate and History: Studies in Honour of Tibor Frank on his 70th birthday*, edited by János Kenyeres, Miklós Lojkó, Tamás Magyarics, and Éva Eszter Szabó, 83–101. Budapest: School of English and American Studies, Faculty of Humanities, Eötvös Loránd University.
- Dabas, Rezső. 1994. “Amerikát okolja a német rendőrállam.” *Magyarság*, February 11, 1994, 1, 6. Retrieved from Arcanum DigiTheca. [https://adt.arcanum.com/hu/view/Magyarsag\\_1994/?query=%22Amerik%C3%A1t%20okolja%20a%20n%C3%A9met%20rend%C5%91r%C3%A1llam%22&pg=88&layout=s](https://adt.arcanum.com/hu/view/Magyarsag_1994/?query=%22Amerik%C3%A1t%20okolja%20a%20n%C3%A9met%20rend%C5%91r%C3%A1llam%22&pg=88&layout=s)
- Devecseri, Tamás. 2007. “Magyar fiúnak nehéz az iskolatáska.” *Amerikai Magyar Népszava – Szabadság*, October 12, 2007, 8. Retrieved from Arcanum DigiTheca. [https://adt.arcanum.com/hu/view/AmerikaiMagyarNepszavaSzabadsag\\_2007\\_2/?query=Magyar%20f%C3%BAnak%20neh%C3%A9z%20az%20iskolat%C3%A1ska&pg=359&layout=s](https://adt.arcanum.com/hu/view/AmerikaiMagyarNepszavaSzabadsag_2007_2/?query=Magyar%20f%C3%BAnak%20neh%C3%A9z%20az%20iskolat%C3%A1ska&pg=359&layout=s)
- Dickinson, William L. 1991. “The Language of Government Act.” *Congressional Record* 137, Part 13, July 18, 1991, House of Representatives, 18829–18881. Retrieved from congress.gov. <https://www.congress.gov/102/crecb/1991/07/18/GPO-CRECB-1991-pt13-5-1.pdf>.

- Fenyvesi, Anna. 2005. "Hungarian in the United States." In *Hungarian Language Contact Outside Hungary: Studies on Hungarian as a Minority Language*, edited by Anna Fenyvesi, 265–318. Amsterdam: John Benjamins. <https://doi.org/10.1075/impact.20.16fen>.
- Gazsó, Dániel. 2020. "Otthon és itthon. A magyar diaszpóra és anyaföldje." PhD diss., Pázmány Péter Catholic University. <https://doi.org/10.15774/PPKE.BTK.2020.006>
- Graedler, Anne-Line. 2014. "Attitudes towards English in Norway: A Corpus-based study of attitudinal expressions in newspaper discourse." *Multilingua* 33 (3–4): 291–312. <https://doi.org/10.1515/multi-2014-0014>.
- Habsburg, Ottó. 1984. "Egy világnyelv felé?" *Chicago és Környéke*, June 23, 1984, 3. Retrieved from Arcanum Digitheca. [https://adt.arcanum.com/hu/view/ChicagoEsKornyeke\\_1984/?query=Egy%20vil%C3%A1gnyelv%20fel%C3%A9%20Habsburg&pg=290&layout=s](https://adt.arcanum.com/hu/view/ChicagoEsKornyeke_1984/?query=Egy%20vil%C3%A1gnyelv%20fel%C3%A9%20Habsburg&pg=290&layout=s)
- Halász, Péter. 1983. "Sauerkraut és társai." *Amerikai Magyar Népszava – Szabadság*, April 15, 1983, 9. Retrieved from Arcanum Digitheca. [https://adt.arcanum.com/hu/view/AmerikaiMagyarNepszava\\_1983\\_1/?query=Sauerkraut%20Hal%C3%A1sz%20P%C3%A9ter&pg=348&layout=s](https://adt.arcanum.com/hu/view/AmerikaiMagyarNepszava_1983_1/?query=Sauerkraut%20Hal%C3%A1sz%20P%C3%A9ter&pg=348&layout=s).
- H.R. 123 EH. 1996. "Bill Emerson English Language Empowerment Act of 1996." 104th Cong., 2nd Sess. <https://www.congress.gov/104/bills/hr123/BILLS-104hr123eh.pdf>.
- H.R. 997 IH. 2023. "English Language Unity Act of 2023." 118th Cong., 1st Sess. <https://www.congress.gov/118/bills/hr997/BILLS-118hr997ih.pdf>.
- H.R. Report 104–723. 1996. "English Language Empowerment Act of 1996." 104th Cong., 2nd Sess., July 30, 1996. <https://www.congress.gov/104/crpt/hrpt723/CRPT-104hrpt723.pdf>.
- Hult, Francis M., and Nancy H. Hornberger. 2016. "Revisiting Orientations in Language Planning: Problem, Right, and Resource as an Analytical Heuristic." *The Bilingual Review/La Revista Bilingüe* 33 (3): 30–49. [https://repository.upenn.edu/gse\\_pubs/476](https://repository.upenn.edu/gse_pubs/476).
- Istook, Ernest. 1996. "English Language Empowerment Act of 1996." *Congressional Record* 142, No. 116. August 1, 1996, House of Representatives, H9738–H9772. Retrieved from congress.gov. <https://www.congress.gov/104/crcr/1996/08/01/142/116/CREC-1996-08-01-pt2-PgH9738-4.pdf>
- Johnson, David Cassels. 2013. *Language Policy*. Basingstoke: Palgrave Macmillan.
- Macías, Reynaldo F. 2000. "Language politics and the sociolinguistic historiography of Spanish in the United States." In *Language in Action: New Studies of Language in Society: Essays in Honor of Roger W. Shuy*, edited by Joy Kreeft Peyton, Peg Griffin, Walt Wolfram and Ralph Fasold, 52–83. Cresskill: Hampton Press.
- Márty, György. 1981. "Szélszövegek: Kétnyelvű analfabéták." *Amerikai Magyar Népszava*, February 27, 1981, 10. Retrieved from Arcanum Digitheca. [https://adt.arcanum.com/hu/view/AmerikaiMagyarNepszava\\_1981\\_1/?query=%22k%C3%A9tnyelv%C5%B1+oktat%C3%A1s%22&pg=177&layout=s](https://adt.arcanum.com/hu/view/AmerikaiMagyarNepszava_1981_1/?query=%22k%C3%A9tnyelv%C5%B1+oktat%C3%A1s%22&pg=177&layout=s).
- Metzger, Nándor. 1974. "Mi az idegen?" *Új Világ*, September 29, 1974, 3. Retrieved from Arcanum Digitheca. [https://adt.arcanum.com/hu/view/UjVilag\\_1974/?query=%22Mi%20az%20idegen%3F%22%20Metzger&pg=446&layout=s](https://adt.arcanum.com/hu/view/UjVilag_1974/?query=%22Mi%20az%20idegen%3F%22%20Metzger&pg=446&layout=s).
- Mikes, György. 1986. "A jenki és a brit." *Szabadság*, July 4, 1986, 12. Retrieved from Arcanum Digitheca. [https://adt.arcanum.com/hu/view/Szabadsag\\_1986\\_07/?query=%20Magyar%20l%C3%A1bnyomok%201986&pg=11&layout=s](https://adt.arcanum.com/hu/view/Szabadsag_1986_07/?query=%20Magyar%20l%C3%A1bnyomok%201986&pg=11&layout=s)
- Mikó Kovács, Zoltán. 1983. "Böngészés." *Californiai Magyarság*, April 15, 1983, 3. Retrieved from Arcanum Digitheca. [https://adt.arcanum.com/hu/view/CaliforniaiMagyarsag\\_1983/?query=%20Californiai%20Magyars%C3%A1g%201983.%20%C3%A1prilis%2015.&pg=142&layout=s](https://adt.arcanum.com/hu/view/CaliforniaiMagyarsag_1983/?query=%20Californiai%20Magyars%C3%A1g%201983.%20%C3%A1prilis%2015.&pg=142&layout=s).
- "Milyen fontos egy szavazat?" 1992. *Magyarság*, May 4, 1992, 3. Retrieved from Arcanum Digitheca. [https://adt.arcanum.com/hu/view/Magyarsag\\_1992/?query=%22Milyen%20fontos%20egy%20szavazat%3F%22&pg=34&layout=s](https://adt.arcanum.com/hu/view/Magyarsag_1992/?query=%22Milyen%20fontos%20egy%20szavazat%3F%22&pg=34&layout=s).
- "Minket nem kényeztettek..." 1982. *Amerikai Magyar Népszava – Szabadság*, January 1, 1982, 2. Retrieved from Arcanum Digitheca. [https://adt.arcanum.com/hu/view/AmerikaiMagyarNepszava\\_1982\\_1/?query=minket%20nem%20k%C3%A9nyeztettek%201982&pg=1&layout=s](https://adt.arcanum.com/hu/view/AmerikaiMagyarNepszava_1982_1/?query=minket%20nem%20k%C3%A9nyeztettek%201982&pg=1&layout=s).

- Nádai, Éva. 1995. "Nyelvháború." *Amerikai Magyar Népszava – Szabadság*, September 29, 1995, 2. Retrieved from Arcanum Digitheca. [https://adt.arcanum.com/hu/view/AmerikaiMagyarNepszavaSzabadsag\\_1995/?query=N%C3%A1dai%20%C3%89va%20nyelv%C3%A1bor%C3%BA&pg=1&layout=s](https://adt.arcanum.com/hu/view/AmerikaiMagyarNepszavaSzabadsag_1995/?query=N%C3%A1dai%20%C3%89va%20nyelv%C3%A1bor%C3%BA&pg=1&layout=s).
- Nánay, Endre. 1981. "Világhelyzet." *Californiai Magyarság*, January 9, 1981, 1–2. Retrieved from Arcanum Digitheca. [https://adt.arcanum.com/hu/view/CaliforniaiMagyarsag\\_1981/?query=t%C3%B6meges%20latin%20%C3%B6z%C3%B6nl%C3%A9s&pg=12&layout=s](https://adt.arcanum.com/hu/view/CaliforniaiMagyarsag_1981/?query=t%C3%B6meges%20latin%20%C3%B6z%C3%B6nl%C3%A9s&pg=12&layout=s).
- . 1985. "Világhelyzet." *Californiai Magyarság*, March 15, 1985, 1–2. Retrieved from Arcanum Digitheca. [https://adt.arcanum.com/hu/view/CaliforniaiMagyarsag\\_1985/?query=N%C3%A1ny%20Vil%C3%A1ghelyzet%20&pg=108&layout=s](https://adt.arcanum.com/hu/view/CaliforniaiMagyarsag_1985/?query=N%C3%A1ny%20Vil%C3%A1ghelyzet%20&pg=108&layout=s).
- "Négyszemközt: Bevándorlás és kétnyelvűség." 1983. *Amerikai Magyar Népszava – Szabadság*, December 23, 1983, 10. Retrieved from Arcanum Digitheca. [https://adt.arcanum.com/hu/view/AmerikaiMagyarNepszava\\_1983\\_2/?query=Ramon%20Santiago%20bev%C3%A1ndorl%C3%A1s&pg=609&layout=s](https://adt.arcanum.com/hu/view/AmerikaiMagyarNepszava_1983_2/?query=Ramon%20Santiago%20bev%C3%A1ndorl%C3%A1s&pg=609&layout=s).
- "Négyszemközt: Bevándorlás és kétnyelvűség 2." 1983. *Amerikai Magyar Népszava – Szabadság*, 1983, December 30, 1983, 10. Retrieved from Arcanum Digitheca. [https://adt.arcanum.com/hu/view/AmerikaiMagyarNepszava\\_1983\\_2/?query=bev%C3%A1ndorl%C3%A1s%20%C3%A9s%20k%C3%A9nyelv%C5%B1s%C3%A9g%20hayakawa&pg=637&layout=s](https://adt.arcanum.com/hu/view/AmerikaiMagyarNepszava_1983_2/?query=bev%C3%A1ndorl%C3%A1s%20%C3%A9s%20k%C3%A9nyelv%C5%B1s%C3%A9g%20hayakawa&pg=637&layout=s).
- ProEnglish. 2023. *Official English Map*. <https://proenglish.org/official-english-map-2/>.
- Pułaczewska, Hanna. 2015. *Student's First Guide to Language Policy and Planning*. Revised ed. Regensburg: Sprachlit.
- Roth, Toby. 1994. "Amendment offered by Mr. Roth." *Congressional Record* 140, Part 4, 1994, House of Representatives, 5567–5801.
- Ruiz, Richard. 1984. "Orientations in language planning." *NABE Journal* 8 (2): 15–34. <https://doi.org/10.1080/08855072.1984.10668464>.
- S. 1109 IS. 2023. "English Language Unity Act of 2023," 118th Cong., 1st Sess.
- Schildkraut, Deborah J. 2003. "American identity and attitudes toward official-English policies." *Political Psychology* 24 (3): 469–99. <http://www.jstor.org/stable/3792322>.
- . 2005. *Press "ONE" for English: Language Policy, Public Opinion, and American Identity*. Princeton: Princeton University Press.
- Schmidt, Ronald, Sr. 2000. *Language Policy and Identity Politics in the United States*. Philadelphia: Temple University Press.
- S.J.Res.72. 1981. "A joint resolution proposing an amendment to the Constitution of the United States with respect to proceedings and documents in the English language," 97th Cong., 1st Session. <https://www.congress.gov/bill/97th-congress/senate-joint-resolution/72?q=%7B%22search%22%3A%5B%22SJRes+72+1981%22%5D%7D&cs=4&r=3>.
- Spolsky, Bernard. 2004. *Language Policy*. Cambridge: Cambridge University Press.
- . 2009. *Language Management*. Cambridge: Cambridge University Press.
- . 2021. *Rethinking Language Policy*. Edinburgh: Edinburgh University Press.
- Ternovszky, Dr. Ferenc. 1983. "Új bevándorlási törvény szükséges!" *Új Világ*, April 22, 1983, 10. Retrieved from Arcanum Digitheca. [https://adt.arcanum.com/hu/view/UjVilag\\_1983/?query=%22k%C3%A9nyelv%C5%B1+oktat%C3%A1s%22&pg=45&layout=s](https://adt.arcanum.com/hu/view/UjVilag_1983/?query=%22k%C3%A9nyelv%C5%B1+oktat%C3%A1s%22&pg=45&layout=s).
- The American Hungarian Federation. 2023a. "About the American Hungarian Federation." <https://www.americanhungarianfederation.org/about.htm>.
- . 2023b. "Intolerance in Slovakia: The Oppressive, Draconian 'Language Law.'" [https://www.americanhungarianfederation.org/news\\_slovakia\\_LanguageLaw.htm](https://www.americanhungarianfederation.org/news_slovakia_LanguageLaw.htm).
- . 2023c. "The Hungarian American Education and Cultural Fund/Amerikai Magyar Oktatási és Kulturális Alap (AMOKA)." [https://www.americanhungarianfederation.org/news\\_AMOKA.htm](https://www.americanhungarianfederation.org/news_AMOKA.htm).
- United States Census Bureau. 2010a. "Language Spoken at Home: Hungarian." *ACS 1-Year Estimates Public Use Microdata Sample, 2010* (LANP) [Data set]. <https://data.census.gov/mdat/#/search?ds=ACSPUMS1Y2010&rv=LANP%28682%29&wt=PWGTP>.

- , 2010b. “Selected Population Profile in the United States.” *ACS 1-year estimates, 2010* (S0201) [Data set]. <https://data.census.gov/table?q=Hungarian&t=Language+Spoken+at+Home&tid=ACSSPP1Y2010.S0201>.
- , 2021a. “Frequently Asked Questions (FAQs) About Language Use.” December 16, 2021. <https://www.census.gov/topics/population/language-use/about/faqs.html>
- , 2021b. “Historical Language Questions.” December 3, 2021. <https://www.census.gov/topics/population/language-use/about/historical.html>
- , 2021c. “Language Spoken at Home: Hungarian.” *ACS 1-Year Estimates Public Use Microdata Sample, 2021* (LANP). [Data set]. <https://data.census.gov/mdat/#/search?ds=ACSPUMS1Y2021&rv=LANP%281582%29&wt=PWGTP>
- , 2021d. “Selected Population Profile in the United States.” *ACS 1-year estimates, 2021* (S0201). [Data set]. <https://data.census.gov/table?q=Hungarian&t=Language+Spoken+at+Home&tid=ACSSPP1Y2021.S0201>
- , 2022. “About the American Community Survey.” June 2, 2022. <https://www.census.gov/programs-surveys/acs/about.html>.
- , 2023. “The Importance of the American Community Survey and the Decennial Census.” <https://www.census.gov/programs-surveys/acs/about/acs-and-census.html>.
- US English. 2022. “About U.S. English. U.S. English Advisory Board.” <https://www.usenglish.org/advisory-board/>.
- Wikipedia. 2023. “Amerikai Magyar Népszava.” Last modified December 18, 2021. [https://hu.wikipedia.org/wiki/Amerikai\\_Magyar\\_N%C3%A9pszava](https://hu.wikipedia.org/wiki/Amerikai_Magyar_N%C3%A9pszava)
- Wiley, Terrence G. 1999. “Comparative historical analysis of U.S. language policy and language planning: Extending the foundations.” In *Sociopolitical Perspectives on Language Policy and Planning in the USA*, edited by Thom Huebner, Kathryn A. Davis and Joseph LoBianco, 17–37. Amsterdam: John Benjamins.
- , 2000. “Continuity and change in the function of language ideologies in the United States.” In *Ideology, Politics and Language Policies: Focus on English*, edited by Thomas Ricento, 67–85. Amsterdam: John Benjamins.
- Wiley, Terrence G., and Haley de Korne. 2014. “Historical orientations to language policy in the United States.” *Language Policy Research Network Brief* (March 2014): 1–4. <http://www.cal.org/lpren/pdfs/briefs/historical-orientations-to-language-policy-in-the-united-states.pdf>.
- Woolard, Kathryn A. 1992. “Language ideology: Issues and approaches.” *Pragmatics* 2 (3): 235–49. <https://doi.org/10.1075/prag.2.3.01woo>.
- Zhou, Minglang. 2015. “Using census data and demography in policy analysis.” In *Research Methods in Language Policy and Planning: A Practical Guide*, edited by Francis M. Hult, and David C. Johnson, 205–16. Malden: Wiley Blackwell.



## On Non-Native Listeners' Ability to Identify Prominence and Pitch Accents in English Monologic Speech

### ABSTRACT

A comparative perceptual study involving two experimental groups with different native languages (Russian and Czech) shows that phonologically trained non-native speakers of English are good at identifying basic suprasegmental features of the English language, namely prominence (sentence stress) and accent types, which potentially makes it possible to use their prosodic annotations when validating cross-language intonation research. The occasional failure of both experimental groups to identify certain accent types is explained in the study by the annotators' mother tongue's prosodic interference: Czech and Russian speakers rely on different acoustic cues when identifying prosodic features in their native languages and transfer this habit to the discrimination of English prosodic characteristics. The study demonstrates that when a prosodic cue is not marked in the speaker's mother tongue, it will likely be ignored in the foreign language.

**Keywords:** prominence, pitch accents, L1 prosodic interference, perceptual analysis

### Sposobnost nerojenih poslušalcev angleščine prepoznavati naglase v angleškem monološkem govoru

#### IZVLEČEK

Primerjalna študija percepcije, ki je vključevala dve eksperimentalni skupini z različnima maternima jezikoma (ruščino in češčino), je pokazala, da so fonološko usposobljeni nerojeni govorci angleškega jezika dobro prepoznali osnovne suprasegmentalne značilnosti angleškega jezika, in sicer prominenco (stavčni poudarek) in vrste naglasa, kar potencialno omogoča uporabo prozodičnih zapisov pri potrjevanju medjezikovnih raziskav intonacije. Občasna neuspešnost obeh eksperimentalnih skupin pri prepoznavanju določenih naglasnih tipov je v študiji pojasnjena s prozodično interferenco maternega jezika testirancev: češki in ruski govorci se pri prepoznavanju prozodičnih značilnosti v svojih maternih jezikih zanašajo na različne akustične signale in to navado prenašajo na razločevanje angleških prozodičnih značilnosti. Študija dokazuje, da če prozodični namig ni označen v maternem jeziku govorca, bo verjetno prezrt tudi v tujem jeziku.

**Ključne besede:** prominenca, višinski poudarki, prozodična interferenca z L1, percepcijska analiza

# 1 Introduction

According to Rietveld and Chen (2006), the established procedure of conducting intonation research requires a mandatory validation of the collected prosodic data by native speakers' perception. However, finding a suitable annotator for conducting perceptual analysis may be complicated, as it means considering a whole range of factors. For instance, the ability of the assessor to understand and annotate a recording has been proven to be significantly influenced by the depth of their previous exposure to a specific language variety or dialect (Mat Daud, Abu Kassim and Mat Daud 2011). Cole, Mahrt and Roy (2017, 322) prove that the most reliable prosodic annotations are made by people "who have strong familiarity with the language variety represented in the materials to be annotated", which means in a time of "global Englishes" it makes sense to look for annotators "locally".

Recruiting auditors "locally" may also be a more rational option as it puts many study parameters under the researcher's immediate control, and allows for more flexibility in experimental designs. However, in non-English speaking countries researchers typically have difficulty recruiting enough phonetically trained native speakers of English to perform the perceptual verification of their results. To solve this problem, Snow et al. (2008) suggest looking into the potential of non-native-speaking annotators.

Although considering this suggestion a reasonable idea, Barry and Andreeva (2001) warn that non-native auditors inevitably attend to the prosodic cues dominant in their native languages, and listeners with native languages that are very different from English may come to unreliable conclusions. In contrast, studies of EFL speakers' suprasegmental features (Rizantseva 2001; Trofimovich and Baker 2006) show that the amount of language contact positively correlates with the learner's ability to recognize English prosodic patterns. McAndrews' (2019) analysis of non-native listeners' identification and discrimination abilities additionally shows that even brief periods of targeted suprasegmental instruction can help minimize foreign language comprehension difficulties and get near-native prosodic annotations.

Snow et al. (2008) believe that if the number of non-native annotators in a research group is big enough, it excludes many perceptual biases and makes the results compatible with the data acquired from native speakers. For instance, Cole, Mahrt and Roy (2017, 306) claim that 10–12 annotators are sufficient in crowd-sourcing tasks to "guarantee the replicability of the annotations at the group level, at least for American English".

Taking the suggestions mentioned above into account, we decided to put non-native listeners' ability to identify certain English suprasegmental features to the test. The present research aims to analyze the degree of correct perception of English prominence and pitch accents by two middle-sized groups of non-native auditors – Czech and Russian (L1) proficient speakers of English – before and after additional phonological training, with the following research questions:

- (1) Do the results of the perceptual analysis of English intonation, namely prominence and accent types, performed by phonologically trained non-native speakers differ significantly from those performed by native speakers of English?

- (2) Do the systemic prosodic differences between the Czech/Russian and English languages interfere with perceiving prominent words and accent types in English monologic speech?
- (3) Is there a meaningful difference in the results of the perceptual analysis of a recorded English text performed by Czech/Russian speakers using different annotation techniques (RPT, ToBI, traditional (British))?
- (4) Can the results of the perceptual analysis performed by non-native speakers be potentially used to validate the intonation of EFL (English as a foreign language) speakers?

## 2 English Prominence Perception and Marking

English prominence perception was chosen as the object of this study following Levis and Silpachai's (2017) research, which considers adequate prominence recognition a key parameter of oral speech perception. Research on prominence (Halliday 1967; Herman 2000; Tench 2011) has proved that phrasal accents in English convey a great deal of structural and discourse information indispensable for effective communication. For instance, the absence of clearly observable prominent words in a dialogue increases the listener's reaction time and deprives the utterances of certain pragmatic functions (Sanderman and Collier 1997).

Another factor influencing our choice was the multifaceted nature of prominence, which, according to Fougeron and Keating (1997), is believed to be a complex phenomenon influenced by both the word's phonetic realization and linguistic contrasts at the same time. Indeed, Cangemi and Grice (2016) show that although the acoustic marking of prominence in English may be rich and multilevel, the same acoustic parameters should be treated differently depending on the word's position in the utterance and its meaningful alignment with the whole text.

Cole, Mo and Hasegawa-Johnson (2010) posit that in most cases the underlying mechanisms of prominence perception are strongly tied to the listeners' expectations, which is illustrated by the fact that in most perceptual tasks of prominence identification the annotators do not rely solely on the acoustic information, but also their expectations of how the utterance will develop. In this way, prominence presents a perfect basis for studying the cumulative effects of various prosodic cues on its perception by groups with different native languages.

Although realizing the complexity of accent structures in English, we suppose that the variety of prominence types in English can be reduced to several intuitively understandable yet valid categories to fit in with our research findings' potential classroom application. Therefore, it is suggested that an intentionally simplified algorithm for annotators with different levels of annotation experience be established to guide them through the basic steps of prominence recognition. This algorithm should include Rapid Prosody Transcription (RPT), traditional (British) prominence marking, and ToBI annotation convention.

*The Rapid Prosody Transcription (RPT)* technique identifies prominent words as the ones "that the speaker has highlighted for the listener to make them stand out" (Cole and Shattuck-

Hufnagel 2016, 8). The annotator has several listenings (passes) of the recording to mark the prosodic boundaries of tone units (the “b-score”) and prominences (the “p-score”) of each word in the recording.<sup>1</sup>

The RPT approach is easy to use, even with annotation novices. It allows sorting focus words irrespective of their actual accent type and other cue variations that make them appear prominent (Example 1).

EXAMPLE 1. RPT annotation of a research sample (abstract) with calculated p-scores.

*|in David **Cameron** (1) |I follow in the **footsteps** (.75) of a great modern Prime **Minister** (1) | under David's **leadership** (1) | the government stabilized the **economy** (1) | reduced the budget **deficit** (1) | and helped **more** (.75) people to work than ever **before** (1) | but David's **true** (1) legacy | is not about the **economy** (1) | but about social **justice** (1) | from the introduction of same-sex **marriage** (1) | to taking people on low **wages** (1) out of **income** (1) tax |* – The annotators marked perceived boundaries (|) and prominences (in bold) in a specially prepared transcript (all punctuation marks and capital letters at the start of the sentences were removed to avoid leading the perception) while listening. The calculated p-scores (in brackets) reflect the perception of the word as prominent in the research group.

The marked boundaries and prominences allow the researcher to divide the identified prominences into “nuclear accented” (preceding the boundary) and “prenuclear accented” (preceding another prominent word) (Example 2).

EXAMPLE 2. Nuclear and prenuclear prominence in RPT annotation.

*|a **language** | becomes a **global** language | because of the **power** of the **people** who **speak** it |* – The utterance contains three tone units separated by perceived boundaries. Prominent words “language” and “global” are the only prominences in the respective units. Each of these words is followed by a boundary (although not immediately in the second case) and classified as a nuclear accent – an indispensable element of a tone unit carrying a contrastive focus. In the third tone group, “speak” is a nuclear accent, and “power” and “people” are prenuclear accents. Prenuclear accents, as a rule, mark new information in the text and play a rhythmical role, balancing the tone group.

Bishop, Kuo and Kim (2020) report that a combination of a distinct pitch movement on a nuclear-accented word with a following boundary tone typically creates a notable pitch change, reportedly accounting for nuclear accents’ easy identification in speech even by non-expert annotators. In contrast, prenuclear tones are more difficult to identify, especially in languages where acoustic marking of sentence prominence, like in Czech, differs from English. Still, Cole and Shattuck-Hufnagel (2016) report that the RPT technique can be a

<sup>1</sup> The b-score and the p-score are calculated as a ratio of annotators who marked the boundary/the prominence in the transcript to the total number of auditors in the study. A figure close to 1 indicates the presence of the respective suprasegmental feature, and a figure close to 0 indicates the absence of the boundary/prominence (Cole and Shattuck-Hufnagel 2016, 9).

reliable means for crowdsourcing large amounts of prosodic information on prominence and boundaries due to its universality.

*Traditional (British) annotation:* Some practical teaching courses still follow O'Connor and Arnold's (1973) traditional (British) annotation system, which presents melody as a directed continuum typically consisting of several stressed syllables in the "head" (at the beginning) of a prosodic phrase followed by a mandatory terminal tone at its end (Example 3).

EXAMPLE 3. A prosodic phrase marked using traditional (British) terms.

because of the → POver of the PEOple who ↘ SPEAK it			
pre-head	+ falling head	+ low fall as a terminal tone	+ tail

Brazil, Coulthard and Johns (1980) stress the importance of the discourse function of the tonic syllable (the terminal tone) in the intonation phrase, which establishes a basic opposition of finality (falling tones) and continuity (rising tones) of the interaction. Similarly, Wells (2006, 15) links the main pitch accent of each intonation phrase with the function of different nuclear tones from the perspective of grammatical or semantic-pragmatic meaning (e.g., "a definitive fall" or an "implicational fall-rise"), thus highlighting the opposition of "falling and non-falling tones" in the nucleus of an intonation phrase.

Concerning the annotation choices, Grabe, Nolan and Farrar (2000) note that at least in four varieties of English, including Standard Southern British English, terminal "falls" and "rises" make up the majority of nuclear tones, which, in our view, makes it possible to limit the inventory of nuclear markings to the fall vs. rise binary opposition. This choice of nuclear tones also seems most relevant for teaching as it reflects the fundamental differences in the syntactic meaning conveyed by nuclear prominence. However, in his recommendations, Brown (2014) emphasizes that nuclear tones do not have a simple one-to-one relationship with meaning, although some general conclusions (e.g., a fall equals finality, whereas a rise means continuity) can still be made.

Summing it up, a tone unit (prosodic phrase) in the traditional (British) terminology always has a terminal tone, viewed as the core of its structure. It may or may not have a preceding head, where the alteration of stressed and unstressed syllables can form different patterns (a "high head", a "falling head", a "stepping head," and others). Traditional (British) annotation identifies prominences as stressed syllables in the "head" if they appear at the beginning of the tone unit and, very generally, subdivides them into falls vs. rises when they appear at its end (as a terminal/nuclear tone).

*ToBI annotation:* The ToBI annotation convention (Pierrehumbert 1980) classifies at least five different pitch accents varying in the placement of the high (H\*) or low (L\*) pitched stressed syllable. Although recognizing the advantages of the ToBI approach for accent identification using visual pitch contours, González-Ferreras et al. (2010) state that the annotators' confusion about the ToBI typology is typically very high in the manual labelling of prominence.

Researchers believe that the ToBI typology works much better – with the agreement rate exceeding 87.7% – only when the annotators are given a binary choice. When presented with a multiclass prominence identification task, even annotation experts report doubts about discriminating different ToBI clusters (González-Ferreras et al. 2010). Therefore, picking the most typical and easily identifiable prenuclear and nuclear accents seems logical to help the annotators limit their confusion.

EXAMPLE 4. A prosodic phrase marked in ToBI convention.

<i>4 because of the <u>power</u> of the <u>people</u> who <u>speak</u> it 4</i>		
H*	H*	H* L-L%
A series of high-pitch accents in a natural downstep followed by a boundary tone.		

As for prenuclear accents in ToBI terms, studies prove that in most standard varieties of English modifications of the H\* tone (mostly L+H\*) come up most frequently (Cho and Keating 2009). González-Ferreras et al. (2010) also state that words with an H\* accent are more strongly associated with perceived prominence among various annotator groups (Example 4), which allows us, following Bishop, Kuo and Kim (2020), to contrast the H\*(L+H\*) prenuclear accent to all other prenuclear types as a binary opposition.

Concerning the nuclear accents, there are again two distinct phrasal tones (L– and H–) and two final boundary tones (L% and H%) at the end of each tone unit, resulting in a range of potential combinations, with H–L% most frequently contrasted with L–L% in four varieties of British English (Grabe, Nolan, and Farrar 2000).

The complete prominence identification algorithm, considering RPT, traditional (British), and ToBI marking, is presented in Figure 1. It reflects the three main approaches to prominence identification presented as binary choices.

The apparent similarity in the approaches consists of dividing the stressed words in the tone unit into two broad categories depending on their closeness to the phrasal boundary. The stressed syllables in the “head” (traditional British annotation terminology) thus correspond to “prenuclear accents” (ToBI terms), and “terminal tones” (traditional (British) annotation terminology) correspond to “nuclear accents” (ToBI terms).

The main differences lie in the types of accents the two systems discriminate. For instance, the British annotation does not assess separate stressed syllables in the head, just marking the resulting contour. In contrast, the ToBI system views each accent separately and even separates a pitch accent and a phrasal/boundary tone, which are viewed as a unit in traditional (British) terms.

In conclusion, we believe that exposing learners to various annotation systems can turn classroom intonation practice (typically aided by traditional (British) terms), crowdsourcing for prosodic data (conducted using the RTP means), and computer-assisted acoustic analysis (recorded via the ToBI system) into a comprehensive FL intonation experience promoting better identification and comprehension of target language prosody.

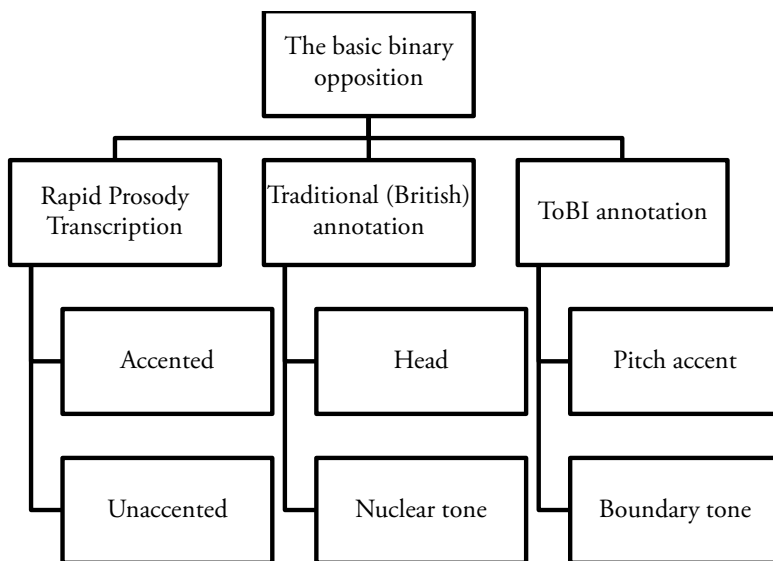


FIGURE 1. The prominence typology across various annotation conventions.

### 3 Cross-Language Differences in Prominence Perception

The perception of many phonetic phenomena in a foreign language relies significantly on their comparative realization in the annotator's/speaker's native language (Vaissière 2005). According to the Markedness Differential Hypothesis (Eckman 1977), it is predicted that all segmental and suprasegmental features of the L2 (foreign language) phonological system can be roughly subdivided into several categories depending on the degree of their markedness in the speaker's L1 (native language) phonological system.

If the marked features of the foreign language phonology differ from the corresponding patterns in the native language, they are also predicted to be more challenging to perceive (Eckman 1977). For instance, Czech learners of English typically have difficulties with English prominence identification, since stress is not a marked prosodic feature in their L1 (Weingartová, Poesová, and Volín 2014).

At the same time, those areas of the target language which are different from the native language but do not form phonological contrasts in L2 will not pose a severe problem to the students, primarily because their impact on the intelligibility of discourse is insignificant. Thus, a different fall/rise ratio in English and Czech does not affect a Czech speaker's speech intelligibility, although it creates an impression of a strong foreign accent (Chamonikolasová 2007).

Since we deal with two groups of non-native speakers, we have to consider the differences in the phonetic realization of prominence in the Russian and Czech languages to point out the L1's potential interference with English prominence perception.

*The characteristics and functions of word stress:*<sup>2</sup> Generally, the nature of word stress in Russian is closer to English stress than Czech stress due to its contrastive nature and the absence of a fixed position in a word. With word stress being a marked feature in their native language, Russian L1 speakers are expected to be more sensitive to English stress placement and its contrastive role in a word or in a sentence than Czech speakers, who primarily rely on the delimitating function of the fixed on the first syllable word stress in their L1 (Skarnitzl and Eriksson 2017).

*The acoustic marking of a stressed syllable:* There is considerable experimental evidence that the prominent constituents of a Russian utterance (i.e., tonal units) are marked acoustically by a perceptible increase in vowel intensity, vowel length change, and higher F0 (Luchkina et al. 2015; Potapova and Potapov 2006). This combination makes Russian word prominence close to word prominence in English, marked by similar acoustic cues (Cruttenden 2014).

In contrast, the fixed on the first syllable of the Czech word stress is not acoustically marked (Skarnitzl and Eriksson 2017). It is usually independent of vowel quality and length, and carries a lower F0 than the following syllables. Skarnitzl and Eriksson thus hypothesize that, unlike in Russian or English, word prominence in Czech is rarely achieved by the stressed syllable prominence, but rather by the acoustic configuration of the whole prosodic word (2017, 3221), with pitch going up and vowels being longer on the post-stressed syllables.

*The acoustic marking of unstressed syllables (vowel reduction):* Another critical characteristic that, in our view, may complicate both the perception and imitation of English prominence by Czech students is the absence of vowel quantitative and qualitative reduction (positional loss of certain characteristics) on unstressed syllables in the Czech language. In contrast to English and Russian, unstressed vowels in the Czech language do not become shorter and do not lose any of their characteristic features in unstressed positions, which, in Weingartová, Poesová and Volín's opinion (2014), explains why Czech listeners demonstrate a weakened perceptual sensitivity to English vowel reductions and experience difficulties identifying distinctions between stressed and unstressed syllables.

*The characteristics of information foci:* Luchkina and Cole (2021, 3) argue that Russian prominence is mostly signal-driven, with positional changes being just an additional means of prominence expression: the acoustic-prosodic expression and the linear order of a sentence, in their opinion, only "reinforce" each other as cues of perceived prominence. In contrast, Skarnitzl and Hledíková (2022) claim that Czech prominence is mainly expectation-driven due to the above-mentioned communicative dynamism (Firbas 1996), which significantly limits the chances of an acoustically marked word being perceived by Czech native speakers as prominent if it comes outside of the expected context.

---

<sup>2</sup> Although we realize that stress at a word level is different from prominence at a sentence level, Cho and Keating (2009, 468) insist that "stress and accent [i.e., sentence stress] represent different degrees along a single scale of prominence and are thus manifested in the same set of physical properties". There is a definite overlap in the use of acoustic dimensions for lexical stress and phrasal accent; therefore, considering the established stress patterns in learners' L1 may contribute to a better understanding of their perception of English prominence.

*The range of terminal tones in L1:* Despite certain apparent similarities in the pitch variation among all three languages, comparative studies of Russian/Czech and English intonation demonstrate a hugely different repertoire of tones employed by the speakers in their native languages, which, due to the L1 intonation transfer, can potentially influence English prominence perception. Leed's (1965) account of Russian intonation states that many tones differ in distribution and meaning from the English tones. In particular, many pretonic rises, non-finite falls, and considerably higher pitch accents of Russian speakers are mentioned as impediments to mutual English-Russian intelligibility. As for the Czech tones, Chamoniolasová (2007) reports smaller melodic variability and a narrower pitch range of Czech speakers, which makes the Czech language sound rather "monotonous" to a foreign ear.

*The boundary markers:* Intonation studies of the Russian language demonstrate that the positional lengthening of the final stressed syllable, also present in English, significantly facilitates the listeners' placement of boundary tones (Leed 1965). In contrast, although Volín and Skarnitzl (2007, 445) prove that in the Czech language "the temporal downtrend in the phrases is, to a large extent, driven by the final deceleration", there is no reliable evidence that Czech speakers regard vowel lengthening as the principal boundary marker as "in languages with fixed stress where every word has stress on the first syllable, it is impossible to separate the correlates of stress from the word boundary effects" (Van Heuven 2018, 25).

To conclude, the comparative analysis of prominence expression means in Russian and Czech shows that the difference in the realization of various stress categories between Russian/Czech and English is significant. The acoustic nature of prominence in Russian seems closer to English prominence, with Czech prominence, on the other hand, being very different from that in English. The absence of salient acoustic markers and unstressed vowel reduction might seriously obstruct English prominence perception by Czech L1 speakers in cases when prominent words do not directly precede a phrasal boundary.

We believe the positive L1 intonation transfer will significantly facilitate the Russian speakers' perception of English prominence. In contrast, Czech speakers of English might be confronted with specific difficulties due to the discrepancy between the stressed syllable's acoustic properties in their L1 and English and the absence of unstressed vowel reduction in the Czech language. On the other hand, meaningful phonological instruction and extensive language exposure are expected to level the differences between the two groups of listeners and make them more sensitive to English phonological contrasts, neutralizing the adverse effects of L1 transfer.

## 4 Method

The annotation experiment was designed as a comparative study involving two groups of participants with different native languages (12 Czech speakers and 14 Russian speakers) but relatively similar in other ways (age, level of English, the study programme). Six British native speakers made up a control group that provided referential data (the so-called "ground truth") for the similarity analysis. All participants signed an informed consent form allowing unrestricted use of their readings in scientific research and publications.

The study was designed as longitudinal research (Figure 2) that involved looking into the two groups' prominence identification over an extended time: at the beginning of the English phonology course (an “amateur ear”, RPT annotation) and 12 weeks later, at the end of it (a “trained ear”, ToBI and traditional (British) annotation).

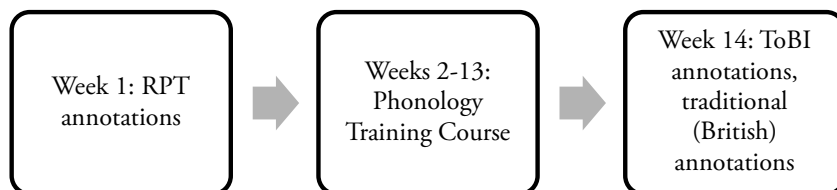


FIGURE 2. Overview of the experiment's structure and timing.

## 4.1 Research Procedure

*Stage 1 (an “amateur ear”):* The first stage of the experiment was meant as a Rapid Prosody Transcription (RPT) annotation of four audio recordings (video footage excluded). The annotators were asked to circle/underline the words that, in their opinion, “stood out” in the speech. Each marked word in the transcript was then attributed p-scores identifying how many annotators marked it as prominent (e.g., [...with a complete8 intolerance0 of failure14...]) and b-scores marking the presence of a perceptual boundary after it. The results of the two annotator groups (Russian and Czech) were compared with the annotations marked by the native-speaker control group (Sub-group 1).

The RPT annotation technique was chosen for the first experimental stage as the most suitable for inexperienced assessors because it requires no special preparation. Annotations were performed after a short introduction of the task. The transcripts for this stage were edited: all punctuation marks and capital letters had been removed from the text to avoid punctuation-guided segmentation bias.

According to Bishop, Kuo and Kim (2020), RPT annotators make more correct prominence identifications when given more passes (listening); however, the number of false identifications also increases proportionately. Since our annotator groups did not have any annotation experience at the beginning of the experiment, they were given three passes of the text: to familiarize themselves with the recording (first pass), to mark the prominences (second pass), and to mark the boundaries (third pass).<sup>3</sup>

*Stage 2 (a “trained ear”):* In the second stage of the experiment, the same groups of annotators had to identify the exact types of nuclear accents (Figure 1). All the participants had the transcripts with the prominent words pre-marked in bold (210 words) to limit the research corpus. The pre-marking was based on the groups' Stage 1 collective performance and included only the words with the highest p-scores (.75–1). The annotators were given two passes of the recordings and were asked to make annotations employing two different annotation

<sup>3</sup> Cole, Mahrt and Roy (2017, 305) report that prominence and boundary marking of the same recording can be performed in any order without any effect on the results.

conventions: ToBI annotations and traditional (British) marks. The texts from Stage 1 were intentionally divided into halves to avoid the annotators' listening to the same recordings four times. While listening, additional short pauses (3–5 sec.) were made by a research assistant to allow some extra time for the annotators to make notes.

The experiment included annotation in ToBI and traditional (British) terms to check whether the groups were still following different annotation traditions after the training course. Following Cole, Mahrt and Roy (2017, 303), in ToBI annotation listeners are often guided by “down-up” pitch evidence from the accompanying pitch/soundwave contour, and a “consensus” annotation using an established inventory provides the best results. In traditional (British) terms, according to Wells (2006, 13), marking tones is more a question of “context and exercise”, with top-down factors such as the utterance's discourse function leading the annotator. The aim of contrasting the two systems was to compare the reliability of the readings and to analyse the prevalence of acoustic vs. contextual factors.

*Statistical analysis:* Both groups' traditional (British) annotation symbols and ToBI marks were then isolated from the transcripts, transferred to digital spreadsheets and contrasted with the control group's annotations with the help of the Excel Compare 2.3 software tool. Cohen's kappa ( $k$ ) was computed to study the similarity of individual annotations with the control group's data. Jaccard coefficients ( $J$ ) were calculated to measure the reliability of small-to-medium groups' annotations of eight sample texts. Finally, the intraclass similarity coefficients (ICC,  $C$ ,  $k$ ) for the two experimental groups (Russian, Czech) and three data sets (RPT, traditional (British), ToBI) were computed to determine the level of internal agreement within each experimental group.

## 4.2 Participants

The Russian annotator group comprised 14 second-year BA students of the Moscow Pedagogical State University (Moscow, Russia) majoring in Anglophone Studies and willing to participate in the experiment. The annotations were made in September and December of the 2022/23 winter term, at the beginning and end of the elective Practical Phonetics 3 course, aimed at exploring and practising English suprasegmentals. All the participants (four males, ten females between the ages of 19 and 22) were advanced speakers of English (C1–C2 level<sup>4</sup>), had a final assessment grade above 74%. The declared level of language proficiency (C1–C2) was guaranteed by a computer-based level test taken by all the participants before the start of the course.

The Czech annotators comprised 12 second-year BA students of the Technical University of Liberec, studying English as one of their two majors. The annotations were conducted within the scope of their SPELB (Experimental Phonology) classes, an elective one-semester course aimed at developing essential analytical and experimental skills in suprasegmental phonology, in September–December 2022. The students of the Czech group (three males, nine females between the ages of 21 and 22) were advanced in English (C1–C2 level) and had successfully passed the final SPELB test (above 70%).

<sup>4</sup> According to the Common European Framework of Reference for Languages (CEFR).

As the description above shows, both experimental groups are perfectly comparable in terms of participants' age, gender composition, academic performance, and language experience, the only meaningful difference between the groups being the students' mother tongue.

### 4.3 The Phonology Training Course

An attempt was made to modify the syllabi of the respective phonetic courses (Practical Phonetics 3/ SPELB (Experimental Phonology)) to minimize the differences between them. Both study courses followed the same syllabus (Appendix 1) and used the same recordings for annotation practice (five texts for traditional (British) annotation and five texts for ToBI practice), amounting to approximately the same amount of instruction time devoted to annotation practice (around four academic hours). The lecturers of both courses were in constant contact to ensure the minimal divergence of the participants' learning experience.

As for the annotation experience, we include in this the extensive suprasegmental training during a 12-week period, including rigorous auditory recognition practice, contrastive differentiation of tones, working with Praat, and targeted annotation practice, sufficient enough to call the learners phonologically trained annotators in the second stage of the experiment (a "trained ear"). The students whose attendance (max. two absences) and academic performance (lower than 70% in the final test) did not meet the established criteria were excluded from the final counts, which explains a slight discrepancy in the number of participants in the experimental groups.

### 4.4 The Control Group

The data for non-native speakers' annotation validation (regarded in this study as the "ground truth") were acquired from the control group of English native speakers. As it was challenging to find a comparable group of English students doing a Phonology course including both traditional (British) and ToBI annotation practice, we compiled the control group from two different sub-groups, emulating the critical parameters of our experimental groups.

The "amateur ear" sub-group (Sub-group 1) comprised four Erasmus students from the University of Reading (UK), who had similar demographics to those of the experimental groups (three females, one male between 21–22 years old). This group comprised students from the BA Primary Education (QTS) programme, and none of the participants had any previous experience in prosody annotation. This "amateur ear" sub-group completely replicated the steps taken by the experimental groups at Stage 1 of the experiment (RPT annotation).

The "trained ear" sub-group (Sub-group 2) consisted of two experienced (PGCE, DipRSA TEFL, LTCL DipTESOL) British lecturers (one male, one female; 40–55 years old) familiar with both annotation systems (traditional (British) and ToBI), meaning that no additional phonological training was needed. Sub-group 2 provided "consensus" ToBI and traditional (British) annotations, meaning that the experts performed the annotations collectively and had an opportunity to discuss their findings.

## 4.5 Materials

Four abstracts of recorded public speeches by four British speakers (two male and two female speakers) were chosen for analysis.<sup>5</sup> The recordings can be characterized as semi-spontaneous monologues and public addresses. Fragments of around two minutes (each up to 209–293 words, and 24–29 grammatical sentences long) were extracted from the video clips and converted into the WAV format. For the second experimental stage, the audio fragments were further divided into halves (each up to one minute long, 95–160 words, and 12–15 grammatical sentences). This was done to present the assessors with different texts by the same speakers to eliminate additional variability.

## 5 Results

### 5.1 Descriptive Statistics of RPT Prominence Identification

In the first stage of the experiment (with an “amateur ear”, RPT marking), a word was considered prominent and included in the final counts if at least 75% of the participants marked it as prominent (its p-score<sup>6</sup> was > .75). The normalized data (Means and Standard Deviations) for all four texts are presented in Table 1.

TABLE 1. The number of identified prominent words (%,<sup>7</sup> raw, standard deviation).

Prominent Words	Text 1 Duration – 1:39 Speaker – JR, female Word count <sup>8</sup> – 218	Text 2 Duration – 1:42 Speaker – DC, male Word count – 293	Text 3 Duration – 1:54 Speaker – TM, female Word count – 258	Text 4 Duration – 1:33 Speaker – JH, male Word count – 209
The Russian annotators	30% (67 words) SD – 8	18% (54 words) SD – 9	24% (61 words) SD – 7	21% (44 words) SD – 6
The Czech annotators	19% (42 words) SD – 11	15% (45 words) SD – 10	18% (47 words) SD – 11	14% (29 words) SD – 10
The control group	32% (70 words) SD – 4	26% (76 words) SD – 2	27% (69 words) SD – 5	28% (58 words) SD – 2

The RPT test’s results demonstrate that both groups of non-native annotators identified fewer prominent words than the native speakers (Sub-group 1). The Russian group identified, on average, 13% fewer prominent words than the control group, and the Czech group marked roughly 40% fewer.<sup>8</sup>

<sup>5</sup> The videos of the annotated recordings are accessible on YouTube:  
 - <https://www.youtube.com/watch?v=wHGqp8lz36c> (3:10–5:03),  
 - <https://www.youtube.com/watch?v=WZI1EjxxXKw> (0:19–1:58),  
 - <https://www.youtube.com/watch?v=FDyZ8trge2E> (1:30–3:37),  
 - <https://www.youtube.com/watch?v=1N2yk0ubNaI> (4:00–6:08).

<sup>6</sup> The p-score was calculated as the ratio of annotators who marked the word as prominent to the total number of annotators in a research group.

<sup>7</sup> Percentage of the total number of words in the respective text.

<sup>8</sup> Here and further in the text, the word count reflects the number of orthographic words, not the phonetic words (feet).

The typological (nuclear-accented words vs. prenuclear-accented words) data comparison of the groups' annotations (Figure 3) revealed that the similarity in the numbers of spotted nuclear accents among all three groups was the highest (10% difference with the control group for the Russian annotators, 24% difference for the Czech annotators).

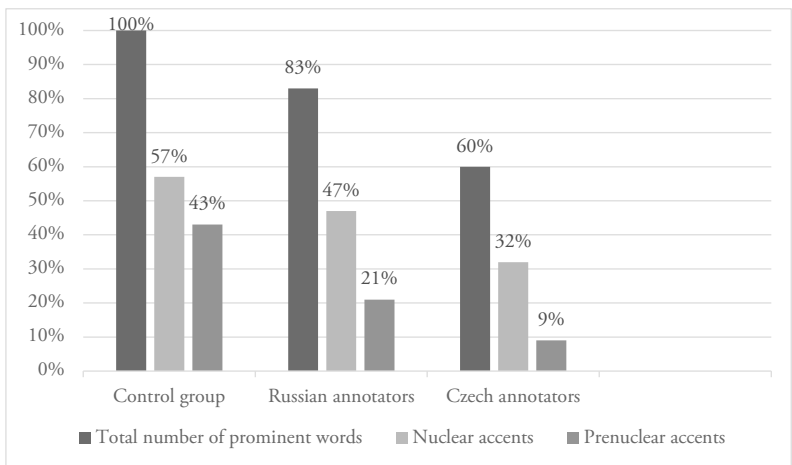


FIGURE 3. The total number and nuclear vs. prenuclear distribution of identified accents (%).

As for the prenuclear accents, the Czech annotators marked almost no prominent words (9%) in this category (34% difference with the control group), while the Russian annotators identified 21% of prenuclear accents (22% difference with the control group). The problems with prenuclear accent identification are illustrated in Example 5.

EXAMPLE 5. Prominence identification (p-scores) by various annotator groups (excerpt).

<i>I cannot (1) remember (1) telling my <b>parents</b> (1)   that I was studying Classics (1)  </i>	The RPT annotation by British native speakers identified four prominent words (two prenuclear and two nuclear prominences) in two tone groups separated by a boundary. The high p-scores of all the words in the sample indicate a high level of annotator agreement.
<i>I cannot (.86) remember (.26) telling my <b>parents</b> (1)   that I was studying Classics (1)  </i>	The RPT annotation by Russian speakers demonstrates a high degree of agreement about the nuclear prominences (“parents” and “classics”) and one prenuclear accent “cannot”. A low p-score for “remember” demonstrates the annotators’ uncertainty.
<i>I cannot (.42) remember (.33) telling my <b>parents</b> (.83)   that I was studying Classics (1)  </i>	The RPT annotation by the Czech speakers shows their agreement only on marking nuclear prominences (“parents” and “Classics”) and weak/no agreement on marking prenuclear accents (“cannot,” “remember”).

This sample from the research corpus provides a typical example of the difficulties faced by non-native listeners in the identification of prenuclear accents. All three groups (British native speakers, Russian listeners and Czech listeners) successfully identified nuclear prominences

(“parents,” “Classics”), whose acoustic marking (a notable pitch change and increased length) is more or less similar in all the three languages and additionally supported by an expectation-driven shift of contextually important words to the end of the utterance.

In contrast, the marking of prenuclear prominences was problematic. The Russian listeners, who acoustically marked “accidental rises” (Leed 1965) in their mother tongue, noticed the acoustic marking of the first prenuclear prominence (“cannot”) but failed to identify the following prenuclear accent (“remember”) since its acoustic marking is expected to be different from the mentioned “accidental rise” due to the natural downstep in F0. The Czech listeners (an “amateur ear”) ignored the acoustic marking of both prenuclear accents probably because they do not have such acoustic cues as high pitch or increased vowel length at the start of the utterance in their mother tongue.

As for individual annotator performances, two annotators in the Czech group (16% of the experimental group) marked considerably fewer prominences (12%) than the other group members, which can be attributed to the assessors’ misunderstanding of the task (e.g., marking only nuclear prominences), rather than their general inability to identify prominent words in English. In contrast, three assessors from the Russian group (20% of the annotators) and one in the Czech group (8%) marked a proportionately (5–7%) bigger number of accents in all four texts than the other assessors in the respective groups, which made their number of identified prominences closer to the control group’s. Still, although with some exceptions, which are reflected in high SD figures for both experimental groups, we believe there is a general trend in non-native prominence identification pointing towards a lower sensitivity to prominence perception, and prenuclear prominence perception in particular.

## 5.2 The Assessment of Individual Annotators’ Performances (ToBI vs. Traditional (British) Marking)

In the second stage, the individual performances of non-native phonetically trained annotators were evaluated with the help of Cohen’s kappa ( $k$ ) by comparing (one-to-one) their annotations with the control group’s readings (Sub-group 2, expert annotators). Cohen’s kappa was then interpreted following McHugh’s (2012) recommendations (Table 2). The extremes of Cohen’s kappa scale are highlighted for the readers’ convenience by *italics* to indicate “none” to “minimal” agreement, and by **bold** to indicate “strong” to “almost perfect” agreement.

TABLE 2. Interpretation of Cohen’s kappa (McHugh 2012).

Value of Kappa ( $\kappa$ )	Level of Agreement	% of Data that are Reliable
0–.20	<i>None</i>	0–4%
.21–.39	<i>Minimal</i>	4–15%
.40–.59	Weak	15–35%
.60–.79	Moderate	35–63%
<b>.80–.90</b>	<b>Strong</b>	64–81%
above .90	Almost perfect	82–100%

Since the annotators had an option not to mark the accent type if they were unsure of its type or they did not have a relevant category in the suggested decision-making algorithm (Figure 1) to reflect the type of accent they heard, the number of accents effectively identified in the second stage (a “trained ear”) was somewhat lower than the total number of prominent words pre-marked for analysis (210 words). The number of words eventually annotated by the participants was 94% (198 words) for the Russian group and 96% (202 words) for the Czech group, with annotators claiming they would need more passes to complete the task.

*The Russian group’s performance results* (Table 3) demonstrate that only one-fourth of the Russian group’s Cohen’s kappa coefficients (in bold) fit into the category of “strong agreement” (with the control group readings), with the mean agreement value for the whole group (.67) carrying “moderate” similarity.

TABLE 3. The Russian group’s Cohen’s kappa ( $\kappa$ ).

The Russian group	ToBI annotation				Traditional (British)			
	Text 1	Text 2	Text 3	Text 4	Text 1	Text 2	Text 3	Text 4
Annot. 1	.67	<b>.81</b>	.57	.59	<b>.80</b>	.77	.69	.71
Annot. 2	.55	.42	.70	.44	.67	.70	.72	.78
Annot. 3	.44	.60	.46	.17	<b>.84</b>	<b>.92</b>	.75	<b>.80</b>
Annot. 4	.24	.30	.19	.28	<b>.84</b>	.72	.67	.72
Annot. 5	.60	<b>.89</b>	.39	.50	.71	.70	.26	.71
Annot. 6	<b>.80</b>	.77	.74	<b>.82</b>	.78	.76	.79	.67
Annot. 6	.68	.68	.77	.78	<b>.94</b>	<b>.81</b>	<b>.82</b>	<b>.87</b>
Annot. 7	.78	<b>.80</b>	.69	<b>.83</b>	.78	<b>.92</b>	<b>.89</b>	<b>.87</b>
Annot. 8	.32	.42	.49	.68	.71	.76	.69	.61
Annot. 9	.67	<b>.81</b>	.77	.59	.64	.72	<b>.80</b>	.71
Annot. 10	.60	.68	.77	<b>.82</b>	<b>.84</b>	.76	.79	.78
Annot. 11	.68	.81	.39	<b>.83</b>	.71	<b>.82</b>	.75	.77
Annot. 12	.32	.70	.57	.50	.77	.76	<b>.89</b>	.77
Annot. 13	.60	<b>.81</b>	.57	.82	.74	<b>.89</b>	<b>.82</b>	<b>.80</b>
Annot. 14	.55	.68	.70	.44	<b>.81</b>	.76	.69	<b>.88</b>

There is an evident trend towards the Russian group’s traditional (British) annotations’ higher reliability (Mean: .77; Median: .78; Mode: .71) than their ToBI markings (Mean: .57; Median: .60; Mode: .60), with nearly one-third of the assessors (28%) providing firm prosodic markings in at least three out of four sample texts. The values for these selected assessors (Annotators 3, 6, 7 and 13; traditional (British) annotation) reach the following peaks: Mean: .84; Median: .83; Mode: .80. The ToBI markings of the Russian group show “moderate” reliability values (Mean: .57; Median: .60; Mode: .60) and contain some “weak” agreement coefficients (7%, in italics).

The irregular distribution of incorrect annotations (highlighted by italics in Table 3), in our view, does not reflect any systemic annotation problems and proves their occasional character, except for Annotator 4, whose low kappa in all four texts (ToBI) most likely signals a misunderstanding of the basic ToBI annotation principles.

McHugh's (2012) interpretation of Cohen's kappa (Table 2) allows us to conclude that the Russian group's traditional (British) markings signal a relatively high degree of data reliability (60%–65% of data is fully reliable). For some annotators (roughly one-third of the group), the level of reliability goes as high as 70% for traditional (British) annotations (bold in Table 4). In contrast, the group's ToBI marks are of worse quality (30%–33% of reliable data).

*The Czech group's performance* (Table 4) reflects an almost even distribution of both high (29%, in bold) and low (14%, in italics) Cohen's coefficients across the annotators in both annotation types. The readings demonstrate slightly higher kappa values of the group's ToBI markings (Mean: .65, Median: .68, Mode: .68) than traditional (British) markings (Mean: .59, Median: .60, Mode: .60). The percentage of unreliable assessments is 14%, which is twice higher than in the Russian group.

TABLE 4. The Czech group's Cohen's kappa ( $\kappa$ ).

The Czech group	ToBI annotation				Traditional (British)			
	Text 1	Text 2	Text 3	Text 4	Text 1	Text 2	Text 3	Text 4
Annot. 1	.78	.56	<b>.84</b>	.70	.61	.74	<i>.31</i>	.62
Annot. 2	.67	.70	.71	.64	<b>.82</b>	<b>.80</b>	.74	<b>.80</b>
Annot. 3	<b>.80</b>	.68	<b>.82</b>	.58	.56	.76	.78	.64
Annot. 4	.74	<b>.88</b>	.72	<b>.93</b>	.46	<b>.80</b>	<b>.81</b>	<b>.86</b>
Annot. 5	<i>.20</i>	.42	.50	.44	.60	.56	.60	.42
Annot. 6	.68	<b>.92</b>	<b>.86</b>	.69	.77	.74	.70	.53
Annot. 6	.68	.58	<i>.34</i>	.54	.56	.40	.51	.60
Annot. 7	.56	.70	.78	<b>.80</b>	.60	.56	<b>.82</b>	<b>.80</b>
Annot. 8	.46	<i>.36</i>	<i>.34</i>	.61	.42	.40	<i>.20</i>	<i>.31</i>
Annot. 9	<b>.82</b>	<b>.92</b>	<b>.80</b>	<b>.88</b>	.71	<b>.80</b>	<b>.88</b>	.64
Annot. 10	.76	.70	.59	.47	.77	.56	.76	.80
Annot. 11	<i>.17</i>	<i>.13</i>	.48	<i>.20</i>	<i>.14</i>	<i>.20</i>	<i>.20</i>	<i>.31</i>
Annot. 12	.72	<b>.86</b>	<b>.81</b>	<b>.87</b>	<b>.82</b>	.76	.81	.94

The character of the invalid annotations' distribution (highlighted by italics in Table 4) among the Czech assessors reflects their non-systematic occurrence, except for Annotators 8 and 11, whose consistently low similarity coefficients in seven out of eight samples may be the result of either limited annotation practice or the annotators' insufficient language exposure, which, as it was mentioned (Weingartová, Poesová, and Volín 2014), can seriously impede Czech students' English prominence perception.

In contrast, some annotators (e.g., Annotators 4, 9, and 12) demonstrated a sustained ability to mark accent types in both annotation systems, with the following averages for the selected group: Mean: .80, Median: .81, and Mode: .80. These annotators' successful performance (as well as the selected reliable annotations of the Russian team), may prove the positive effects of preliminary phonological training on English prominence perception. The successful annotations can also be attributed to the students' high English proficiency (C1–C2).

In conclusion, the recommended quantitative interpretation of Cohen's kappa (McHugh, 2012) signals the moderate reliability of the Czech group's ToBI annotations (48%–55% of reliable data) and weak reliability of their traditional (British) assessments (30–35%).

### 5.3 The Assessment of Groups' Collective Performances: Jaccard Similarity Indices

To assess the collective reliability of 12/14-member teams made up of non-native phonologically trained annotators, Jaccard similarity indices ( $J$ ) were computed for eight texts (four were marked in ToBI and four using traditional (British) annotations). The experimental group's most common annotation mark for each of the pre-marked 210 words was compared to the respective control group's marking.

The computed Jaccard similarity indices for both annotator groups (Figures 4 and 5) were evaluated following recommendations for the similarity coefficients' evaluation (Verma and Aggarwal 2020), with values of 1%–20% indicating “none” to “slight” similarity, 21%–40% – “fair,” 41%–60% – “moderate,” 61%–80% – “substantial,” and 81%–100% – “almost perfect” similarity.

The results of the Russian group's comparison (Figure 4) demonstrate “moderate” to “substantial” levels of annotation similarity with the native speakers' data, with a maximum of 80% achieved in one of the four texts annotated with the help of traditional (British) annotation symbols. The ToBI annotations (ranging from 45% to 65%) in three out of the four ToBI marked texts demonstrate a slightly lower level of annotation similarity than do the traditional (British) annotations (65% to 80%), which is quite in line with the group's Cohen's kappa values.

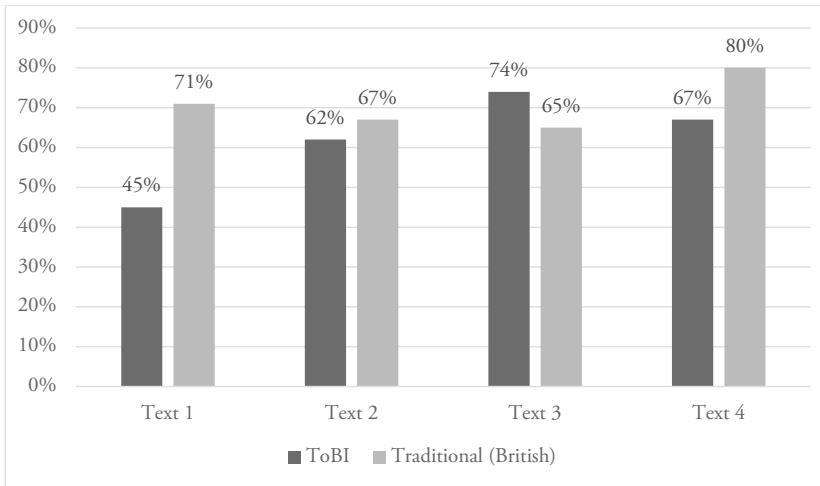


FIGURE 4. The Jaccard ( $J$ ) similarity indices for the Russian annotations.

The Czech group's Jaccard indices (Figure 4) demonstrate “moderate” to “substantial” similarity with the control group's marking. The experimental group had a lower similarity

in the annotations of two texts performed using traditional (British) symbols (43%–63%). At the same time, the ToBI annotations done by the Czech assessors showed a consistently better correlation with the native speakers' data in all four texts (65%–82%). These findings also support the evidence of the group's better ToBI performance calculated with the help of Cohen's kappa.

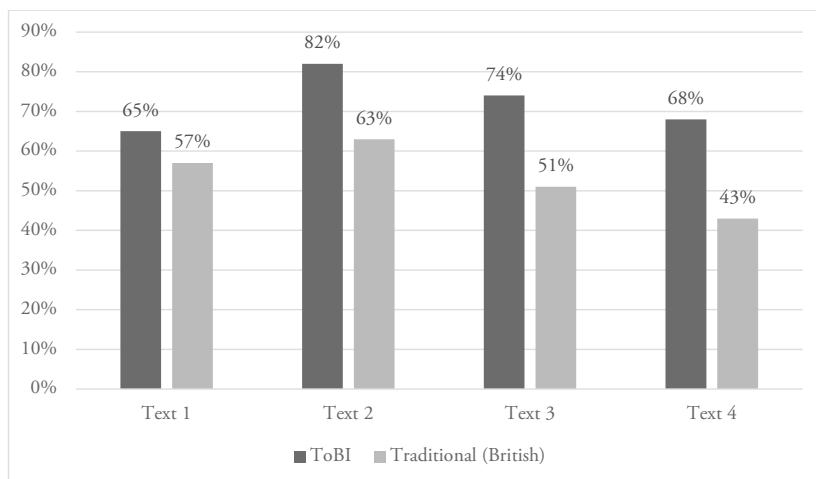


FIGURE 5. The Jaccard ( $J$ ) similarity indices for the Czech annotations.

#### 5.4 Intragroup Reliability Coefficients

Finally, a two-way, consistency, average-measure intraclass correlation coefficient (ICC ( $C, k$ ), with  $k$  standing for the number of participants in the group) was computed to examine the internal agreement within two experimental groups of assessors. The number of identical assessments of a unit (an accent type) across the group in tree data sets (RPT, ToBI, traditional (British)) was referred to the total number of assessments of a particular unit (word) and statistically normalized by the number of group members (Table 5).

Table 5. ICC ( $C, k$ ) reliability coefficients.

	Stage 1 (RPT annotation)	Stage 2 (traditional (British) marking)	Stage 2 (ToBI annotation)
<b>The Russian group, <math>k=14</math></b>	.88	.71	.68
<b>The Czech group, <math>k=12</math></b>	.76	.62	.70

Under Cicchetti's (1994, 286) benchmarks for the ICC evaluation ( $< .40$  – “poor,”  $.40$ – $.59$  – “fair,”  $.60$ – $.75$  – “good,” and  $> .75$  – “excellent significance”), both groups demonstrate “good” to “excellent” homogeneity of their annotations.

The groups' RPT annotations, naturally, have higher intraclass correlation indices ( $.76$  – the Czech group,  $.88$  – the Russian group), as the annotators had a limited choice of options

(“stressed” vs. “unstressed” word). At the same time, since the number of options doubled at the second stage (an “experienced” ear), the intragroup agreement coefficients went slightly down (min.= .62, max.= .71), while still remaining high enough to substantiate the validity of the groups’ results.

## 6 Discussion

### 6.1 The Performance of the Russian and Czech Experimental Groups vs. the Control Group

The results of the two-stage perceptual experiment demonstrate that Russian and Czech speakers are generally less sensitive to prominence perception in English monologues than native speakers of English. It is assumed that the fact that both experimental groups identified at least 20% fewer prominent words and had 7–14% unreliable readings can be explained by either their insufficient phonological training (Stage 1) or some cumulative effects of the prosodic differences between the speakers’ native languages and English that interfere with English prominence perception.

In line with earlier experiments (Bishop, Kuo, and Kim 2020), which prove that annotators identify specific types of prominence more easily than others, our findings demonstrate that nuclear-accented words are generally more noticeable for non-native annotators than prenuclear prominences. We believe that the relative similarity of terminal pitch movements on nuclear accents, the terminal tones’ delimitating function, and the similarity of their pragmatic load in all three languages lead to their better recognition.

In contrast, prenuclear accents in non-native speaker annotations demonstrate significantly lower recognition indices, probably because they carry primarily acoustic (local/signal-driven) cues, which are believed to be more difficult to perceive for non-native speakers than expectation-driven semi-global and global factors. The Russian annotators did not recognize one-fifth of prenuclear accents, and the Czech group failed to identify one-third. The results of the Russian group are believed to be better because of the similarity of certain local acoustic features in Russian and English syllables, facilitating prenuclear accent identification.

### 6.2 The Czech/Russian Languages’ Prosodic Interference in English Prominence Perception

We believe that the closeness of the acoustic marking of English and Russian accented words, as shown in some respects in earlier research – such as peak pitch values, noticeable stressed vowel’s positional lengthening, and higher loudness on the focus words – contributed to the Russian group’s better English prominence recognition, accounting for a higher number of successfully annotated words in both experimental stages. The reported wider pitch range inherent in Russian intonation contours (another marked prosodic feature in Russian) may also assist in the Russian annotators’ better recognition of English accents. Additionally, phonological contrasts that result from the reduction of unstressed vowels, present in English and Russian, are generally thought to facilitate the Russian group’s accent recognition.

In contrast, the Czech annotators, who, according to the aggregated data, have neither peak pitch values nor positional vowel lengthening on the stressed syllables in their L1, found English prominence perception a bit more challenging. In our opinion the main difficulty, which may generally serve as an impediment to more successful prominence identification for the Czech speakers of English, lies in the absence of vowel reduction in unstressed syllables as a marked feature in the Czech language. This accounts for the Czech speakers' lower sensitivity to vowel-length contrasts, although it is a valuable acoustic cue for English native speakers. The narrower pitch range in the Czech language reported in earlier studies may also significantly complicate the perception of English prenuclear accents at the beginning and, particularly, in the middle of the intonation phrase, with high pitch, although significant for the native speakers of English in marking the prenuclear accents, serving as just an irrelevant distractor for the Czech speakers.

### 6.3 The Effects of Using Different Annotation Techniques

The decision to compare and contrast different methods of prominence marking in the present study was made because earlier comparative studies of annotation types in the phonology-teaching context showed no significant difference in their classroom application, and left the choice of annotation system to the theoretical preferences of the teacher (Toivanen 2005). However, we believe that students' L1 may also influence the researcher's potential choice of annotation technique. The chosen annotation method may better reflect the non-native annotators' dominant L1 prosodic cues and, therefore, may provide higher annotation precision in the target language.

The use of RPT brought highly reliable results with untrained assessors. In line with earlier studies (Cole and Shattuck-Hufnagel 2016), we thus recommend its wider application with annotation novices, mainly because RPT does not require much preparatory practice.

As for ToBI and traditional (British) annotation methods, the acquired readings show mixed results. Noticeable discrepancies in the ToBI vs. traditional (British) annotations' similarity indices demonstrate the Russian assessors' better performance with the traditional (British) annotation technique, which, as expected, better reflects "top-down" contextual cues. In contrast, the Czech group had consistently better results in the ToBI marking, likely due to the instilled practice of "down-up" acoustic cue perception. Nevertheless, it seems too early to draw conclusions about the students' preferences for one or the other system, as more factors, such as their annotation experience and the quality of annotation teaching, should also be considered.

### 6.4 Non-native Speakers' Potential in EFL Intonation Research Validation

Our results demonstrate that the experimental groups' annotations can be characterized by "moderate" (60%) to "substantial" (70 %) similarity with native speakers' assessments, with the number of incorrect annotations (7%–14%) being only slightly higher than the results seen in earlier studies (80% reliability, 5% of incorrect assessments, for the Russian assessors in Potapova and Potapov (2006)).

The computed intraclass correlation coefficients show "good" to "excellent" data homogeneity of the 12/14-member groups in all three types of tasks (RPT, traditional (British), ToBI).

These alignment coefficients are in line with earlier findings (Cole, Mahrt, and Roy 2017), which means that attracting medium-sized annotator groups to data validation is reasonably practical.

Attracting 12 to 14 annotators to data validation allows the researcher to overcome the effects of individual perceptual differences and makes it possible for small to medium-sized groups of non-native speakers to be employed for cross-language validation of prominent words in various experimental studies involving EFL speakers. However, it should be noted that prior phonological training, a general level of foreign language proficiency, and a suitable annotation method, should also be considered when forming a non-native annotation team.

## 7 Limitations

The present research had the following limitations that were addressed with great care in the present paper and should be addressed in further studies.

The experimental design of annotation tasks needs to consider the issue of the annotators' English language exposure and annotation experience, as closer attention to these parameters may provide explanations for several annotation mistakes. Our study addressed the issue of language proficiency by conducting a computer-based level test in both research groups. Students' attendance scores also served as a criterion for their participation in the final research stage.

Giving annotators more time to think (while keeping the same number of passes), excluding pre-marked words from the transcripts, and departing from the pre-set algorithm of prominence marking will potentially increase the variability of identified accents and prosodic cues the listeners rely on. On the other hand, using binary choices as we did in our study (the suggested binary-choice algorithm in Figure 1) ensures the homogeneity of research standards for the experimental groups and the materials to be annotated.

Finally, the occasionally mixed record of ToBI vs. traditional (British) annotations evident from our study calls for more attention to the choice of annotation technique in experimental papers. Despite the apparent preference of certain pronunciation instructors for the use of traditional (British) annotation principles, the vast majority of current intonation research uses the ToBI annotation method, and students should have enough in-class annotation practice to be able to confidently use this system in their research papers.

## 8 Conclusion

The statistical data acquired in this study, namely the sufficient similarity and reliability coefficients, demonstrate that language proficient and phonologically trained non-native speakers of English can effectively recognize and identify word prominence and basic accent types in English monological speech. As suggested in earlier papers (McAndrews 2019; Snow et al. (2008)), the results of phonologically trained non-native listeners' perceptual analysis, although with certain reservations, can be successfully used to validate English intonation research in non-English speaking countries where English is predominantly spoken as a second or foreign language.

The study results demonstrate that trained non-native auditors can use various annotation techniques, such as Rapid Prosody Transcription, ToBI marking, and traditional (British) annotation, to mark key prosodic phenomena in English. Although neither annotator group in our study expressed a clear preference for one or another annotation system, it can be assumed that the choice of annotation technique for educational or research purposes should be linked to the annotator's prior phonological training.

Our findings also show that positive intonation transfer (i.e., easier recognition of similar, often marked features) may contribute to a better perception of the prosodic cues similar in the annotators' native (Russian and Czech) and target (English) languages. Despite the negative L1 interference, evident from the annotators' more problematic recognition of the prosodic features unmarked in their native languages (e.g., prenuclear prominence), the experimental groups of Czech and Russian speakers still demonstrate a sustained ability to mark English prominence, which, presumably, can be further improved through additional language and annotation practice.

In conclusion, we believe that more cross-language perceptual studies involving annotators with different native languages are needed to specify the cumulative role of various prosodic cues in English learners' intonation recognition, as their results can help EFL teachers focus prosodic instruction on the aspects crucial for language intelligibility.

## Acknowledgements

The author expresses profound gratitude to the colleagues who helped collect data and conduct the present experiment.

## References

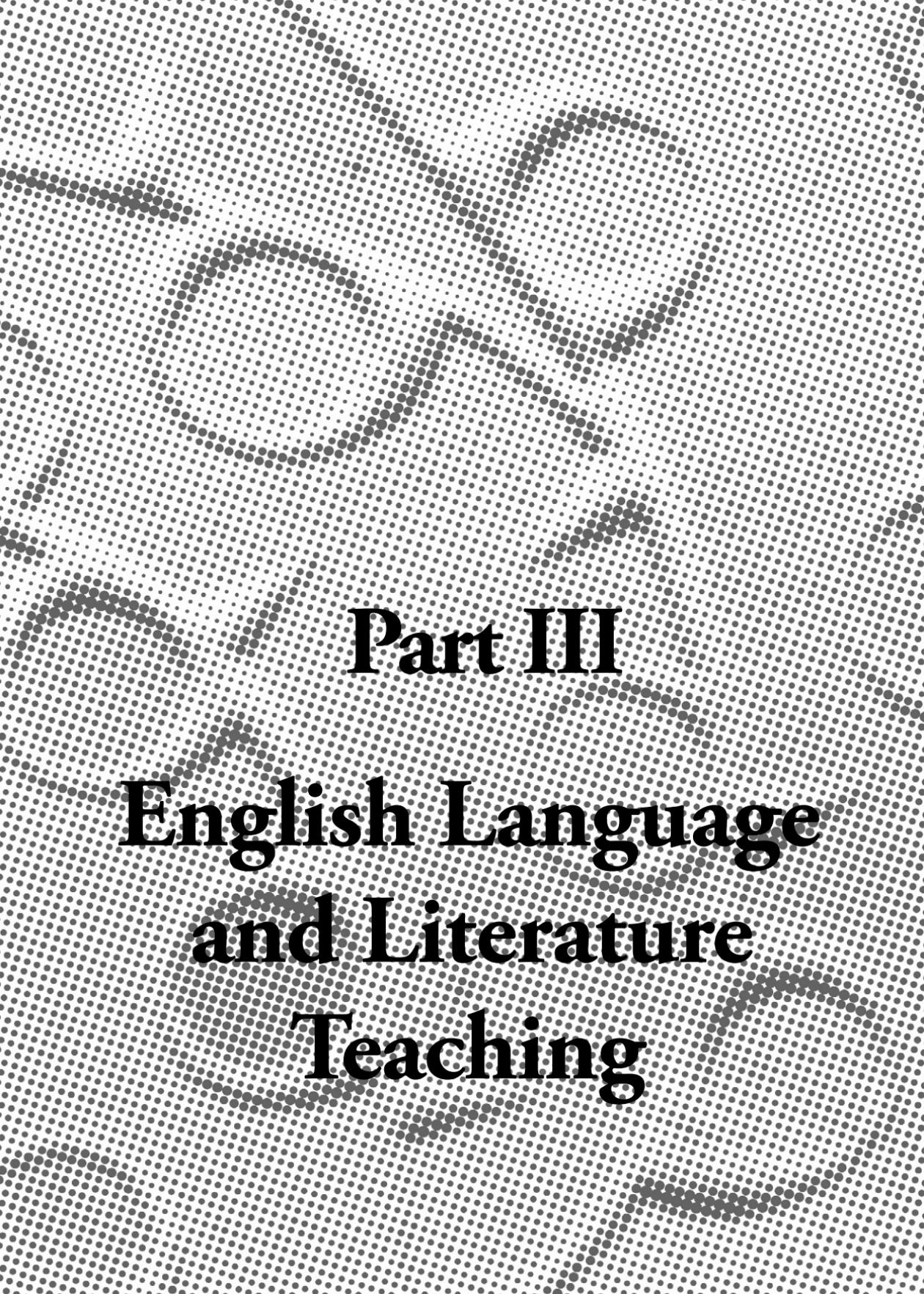
- Barry, William, and Bistra Andreeva. 2001. "Cross-language similarities and differences in spontaneous speech patterns." *Journal of the International Phonetic Association* 31 (1): 51–66. <https://doi.org/10.1017/S0025100301001050>.
- Bishop, Jason, Grace Kuo, and Boram Kim. 2020. "Phonology, phonetics, and signal-extrinsic factors in the perception of prosodic prominence: Evidence from rapid prosody transcription." *Journal of Phonetics* 82. <https://doi.org/10.1016/j.wocn.2020.100977>.
- Brazil, David, Michael Coulthard, and Christopher Johns. 1980. *Discourse Intonation and Language Teaching*. London: Longman.
- Brown, Adam. 2014. *Pronunciation and Phonetics: A Practical Guide for English Language Teachers*. London: Routledge.
- Cangemi, Francesco, and Mary E. Grice. 2016. "The importance of a distributional approach to categoricity in autosegmental metrical accounts of intonation." *Laboratory Phonology* 7 (1): 9. <https://doi.org/10.5334/labphon.28>.
- Chamonikolasová, Jana. 2007. *Intonation in English and Czech Dialogues*. Brno: Masarykova Univerzita.
- Cho, Taehong, and Patricia Keating. 2009. "Effects of initial position versus prominence in English." *Journal of Phonetics* 37: 466–85. <https://doi.org/10.1016/j.wocn.2009.08.001>.
- Cicchetti, Dante. 1994. "Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology." *Psychological Assessment* 6: 284–90. <https://doi.org/10.1037/1040-3590.6.4.284>.
- Cole, Jennifer, Timothy Mahrt, and Joseph Roy. 2017. "Crowd-sourcing prosodic annotation." *Computer Speech & Language* 45: 300–25. <https://doi.org/10.1016/j.csl.2017.02.008>.

- Cole, Jennifer, Yoonsook Mo, and Mark Hasegawa-Johnson. 2010. "Signal-based and expectation-based factors in the perception of prosodic prominence." *Laboratory Phonology* 1: 425–52. <https://doi.org/10.1515/labphon.2010.022>.
- Cole, Jennifer, and Stefanie Shattuck-Hufnagel. 2016. "New methods for prosodic transcription: Capturing variability as a source of information." *Laboratory Phonology* 7 (1): [8]. <https://doi.org/10.5334/labphon.29>.
- Cruttenden, Alan. 2014. *Gimson's Pronunciation of English*. London: Routledge.
- Eckman, Fred. 1977. "Markedness and the contrastive analysis hypothesis." *Language Learning* 27 (2): 315–30. <https://doi.org/10.1111/j.1467-1770.1977.tb00124.x>.
- Firbas, Jan. 1996. "Mobility of clause constituents and functional sentence perspective." In *Discourse and Meaning: Papers in Honour of Eva Hajičová*, edited by Barbara H. Partee and Petr Sgall, 221–33. Amsterdam: John Benjamins.
- Fougeron, Cécile, and Patricia Keating. 1997. "Articulatory strengthening at edges of prosodic domains." *The Journal of the Acoustical Society of America* 101 (6): 3728–40. <https://doi.org/10.1121/1.418332>.
- González-Ferreras, Clara, Carmen Vivaracho-Pascual, Daniel Escudero-Mancebo, and Victoria Cardenoso-Payo. 2010. "On the automatic ToBI accent type identification from data." *Interspeech 2010*: 142–45. <https://doi.org/10.21437/Interspeech.2010-70>.
- Grabe, Esther, Francis Nolan, and Kari Farrar. 2000. "Pitch accent realization in four varieties of British English." *Journal of Phonetics* 28 (2): 161–85. <https://doi.org/10.1006/jpho.2000.0111>.
- Halliday, M. A. 1967. *Intonation and Grammar in British English*. The Hague: Mouton.
- Herman, Robert. 2000. "Phonetic markers of global discourse structures in English." *Journal of Phonetics* 28: 466–93. <https://doi.org/10.1006/jpho.2000.0127>.
- Hirschberg, Julia, and Mary Beckman. 1994. "The ToBI annotation conventions." Retrieved from <http://www.cs.columbia.edu/~julia/files/conv.pdf>.
- Leed, Robert L. 1965. "A contrastive analysis of Russian and English intonation contours." *The Slavic and East European Journal* 9 (1): 62–75. <https://doi.org/10.2307/304390>.
- Levis, John M., and Alif O. Silpachai. 2017. "Prominence and information structure in pronunciation teaching materials." *Pronunciation in Second Language Learning and Teaching Proceedings* 9 (1): 216–29.
- Luchkina, Tatiana, and Jennifer Cole. 2021. "Perception of word-level prominence in free word order language discourse." *Language and Speech* 64 (2): 381–412. <https://doi.org/10.1177/0023830919884089>.
- Luchkina, Tatiana, Varun Puri, Pavithra Jyothi, and Jennifer Cole. 2015. "Prosodic and structural correlates of perceived prominence in Russian and Hindi." In *Proceedings of the 18th International Congress of Phonetic Sciences*. [https://lbp-us-e1.wpmucdn.com/sites.northwestern.edu/dist/f/1603/files/2017/01/Luchkina-Jyothi-Sharma-Cole\\_ICPHS\\_2015-2n15ibp.pdf](https://lbp-us-e1.wpmucdn.com/sites.northwestern.edu/dist/f/1603/files/2017/01/Luchkina-Jyothi-Sharma-Cole_ICPHS_2015-2n15ibp.pdf).
- Mat Daud, Nuraihan, Noor Lide Abu Kassim, and Nor Shidrah Mat Daud. 2011. "Students as assessors." In *Language Studies in the Muslim World*, 1–18. IIUM Press. <https://doi.org/10.13140/2.1.2804.4162>.
- McAndrews, Maura. 2019. "Short periods of instruction improve learners' phonological categories for L2 suprasegmental features." *System* 82: 151–60. <https://doi.org/10.1016/j.system.2019.04.007>.
- McHugh, Mary Louise. 2012. "Interrater reliability: The Kappa statistic." *Biochemia Medica* 22 (3): 276–82.
- O'Connor, Joseph, and Gimson Arnold. 1973. *Intonation of Colloquial English*. Bristol: Longman.
- Pierrehumbert, Janet. 1980. "The Phonology and Phonetics of English Intonation." Ph.D. diss., Massachusetts Institute of Technology.
- Potapova, Raisa, and Vsevolod Potapov. 2006. *Jazyk, Rech, Lichnost*. Moscow: Jazyki Slavianskoy Kultury.
- Riazantseva, A. 2001. "Second language proficiency and pausing. A study of Russian speakers of English." *Studies in Second Language Acquisition* 23 (4): 497–526. <https://doi.org/10.1017/S027226310100403X>.
- Rietveld, Toni, and Alan Chen. 2006. "How to obtain and process perceptual judgments of intonational meaning." In *Methods in Empirical Prosody Research*, edited by Stefan Sudhoff, Denisa Lenertova, Roland Meyer, Sandra Pappert, Petra Augurzy, Ina Mleinek, Nicole Richter and Johannes Schließer, 283–319. Berlin: Mouton de Gruyter. <https://doi.org/10.1515/9783110914641>.

- Sanderman, Ann, and Robert Collier. 1997. "Prosodic phrasing and comprehension." *Language and Speech* 40 (4): 391–409. <https://doi.org/10.1177/002383099704000405>.
- Skarnitzl, Radek, and Anders Eriksson. 2017. "The acoustics of word stress in Czech as a function of speaking style." *Interspeech 2017*: 3221–25. <https://doi.org/10.21437/Interspeech.2017-417>.
- Skarnitzl, Radek, and Hana Hledíková. 2022. "Prosodic phrasing of good speakers in English and Czech." *Frontiers in Psychology* 13: 1–13. <https://doi.org/10.3389/fpsyg.2022.857647>.
- Snow, Rion, Brendan O'Connor, Dan Jurafsky, and Andrew Ng. 2008. "Cheap and fast – but is it good? Evaluating non-expert annotations for natural language tasks." In *Proceedings of the 2008 Conference on Empirical Methods in Natural Language Processing*, edited by Mirella Lapata and Hwee Tou Ng, 254–63. Honolulu: Association for Computational Linguistics.
- Tench, Paul. 2011. *Transcribing the Sound of English: A Phonetics Workbook for Words and Discourse*. Cambridge: Cambridge University Press.
- Toivanen, Juhani. 2005. "ToBI or not ToBI? Testing two models in teaching English intonation to Finns." In *Proceedings of Phonetics Teaching and Learning Conference*, 2005. <https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=4f7274b4fab041bde8d81f2307b93418737e9ddb>.
- Trofimovich, Pavel, and William Baker. 2006. "Learning second language suprasegmentals: Effect of L2 experience on prosody and fluency characteristics of L2 speech." *Studies in Second Language Acquisition* 28 (1): 1–30. <https://doi.org/10.1017/S0272263106060013>.
- Vaissière, Jacqueline. 2005. "Perception of intonation." In *The Handbook of Speech Perception*, edited by David Pisoni and Robert Remez, 236–63. Malden: Blackwell Publishing.
- Van Heuven, Vincent J. 2018. "Acoustic correlates and perceptual cues of word and sentence stress." In *The Study of Word Stress and Accent: Theories, Methods, and Data*, edited by René Goedemans, John Heinz, and Harry van der Hulst, 15–59. Cambridge: Cambridge University Press.
- Verma, Vishesh, and Ritu K. Aggarwal. 2020. "A comparative analysis of similarity measures akin to the Jaccard index in collaborative recommendations: An empirical and theoretical perspective." *Social Network Analysis and Mining* 10: 43. <https://doi.org/10.1007/s13278-020-00660-9>.
- Volín, Jan, and Radek Skarnitzl. 2007. "Temporal downtrends in Czech read speech." *Interspeech 2007*: 442–45. <https://doi.org/10.21437/Interspeech.2007-214>.
- Weingartová, Lucie, Katarína Poesová, and Jan Volín. 2014. "Prominence contrasts in Czech English as a predictor of learner's proficiency." *Speech Prosody* 7: 236–40. <https://doi.org/10.21437/SpeechProsody.2014-35>.
- Wells, John Christopher. 2006. *English Intonation: An Introduction*. Cambridge: Cambridge University Press.

## Appendix 1. The Phonology Course Syllabus

Study Week	The topics discussed:
Week 1	Introduction to experimental phonology. Methods of phonetic and phonological analysis. Phonetic and phonological mistakes. Acoustic and auditory (perceptual) analysis of speech.
Week 2	Components of pronunciation. Suprasegmentals. Syllabic structure of a language. Falling and rising tones. Fall/rise perception and production practice.
Week 3	Stress. Word stress. Functions of word stress. Acoustic properties of stress. Stress placement practice.
Week 4	Stress in English and in Czech/Russian. Perceptual and acoustic differences. Stress identification practice.
Week 5	Prosody of speech. Intonation and prosody. Components of intonation. The structure of an utterance. Pitch and pitch accents. Pauses and breaks.
Week 6	Functions of intonation. Linking tones to functions. Phonological analysis of an utterance. Annotation practice.
Week 7	Prominence. Functions of prominence. Focus words. Informational structure of an utterance. Acoustic correlates of prominence. RPT. Annotation practice.
Week 8	Traditional (British) annotation. Speech melody. Rhythm. Differences between Czech/Russian and English intonation. Annotation practice.
Week 9	Phonostylistics. Loudness and tempo variations. Connected speech processes. Unstressed vowel reduction. Annotation practice.
Week 10	Introduction to the computer analysis of speech. Praat analysis. Acoustic correlates of perceptual parameters. Annotation practice.
Week 11	Autosegmental phonology. Introduction of ToBI. Pitch accents and boundary tones. Annotation practice.
Week 12	Regional intonation variation. ToBI annotation practice. Comparison of ToBI and traditional (British) system. Annotation practice.



**Part III**

**English Language  
and Literature  
Teaching**



Nikola Dobrić  
University of Klagenfurt, Austria

2024, Vol. 21 (1), 91-110(316)  
journals.uni-lj.si/elope  
<https://doi.org/10.4312/elope.21.1.91-110>  
UDC: 808.1:37.091.279.7

## Assessment-Relevant Stimuli and Judging of Writing Performances – From Micro-Judgments to Macro-Judgments

### ABSTRACT

The contemporary practice of rating writing performances is grounded in an approach known as judging, which is done to avoid paying conscious attention to discrete elements in texts. Instead, it involves accounting for the overall impression made by a writing performance. However, studies have indicated that while this may be true on a conscious level, the concrete stimuli in texts still preconsciously influence the forming of such overall impressions. What is left largely unnoticed is that most assessment-relevant stimuli require the use of judging to be perceived as such. This implies that an overall, macro-judgment of a writing performance (expressed normally as a score) comprises individual (and largely preconsciously generated) micro-judgments coming together into a complex and non-linear combination-count. The paper presents an argument in favour of such a composition of judgments, demonstrates it empirically by means of a case study, and then discusses the wider consequences of this different perspective on judging.

**Keywords:** writing assessment, judging, macro-judgment, micro-judgment, assessment-relevant stimuli

## Dražljaji, pomembni za ocenjevanje, in presoja pisnih izdelkov – od mikropresoj do makropresoj

### IZVLEČEK

Sodobna praksa ocenjevanja pisnih izdelkov temelji na pristopu, znanem kot presoja. Pri njem se izogibamo zavestnemu osredinjanju na posamezne prvine besedila in namesto tega ocena temelji na splošnem vtisu o besedilu. Vendar so študije pokazale, da to sicer velja na zavedni ravni, nezavedno pa na oblikovanje splošnega vtisa še vedno vplivajo konkretni dražljaji v besedilih. Praviloma se ne zavedamo, da večina dražljajev, pomembnih za ocenjevanje, sama zahteva presojo, da bi jih prepoznali kot take. To pomeni, da je splošna presoja pisnega izdelka na makroravni (običajno v obliki ocene) sestavljena iz posameznih (in v veliki meri nezavedno ustvarjenih) mikropresoj, ki se povežejo v zapleteno in nelinearno preštevanje kombinacij. V prispevku je predstavljen argument v prid takšni sestavi presoj, empirično podprt s študijo primera, v nadaljevanju pa so obravnavane širše razsežnosti tovrstnega drugačnega pogleda na ocenjevanje pisnih izdelkov.

**Ključne besede:** ocenjevanje pisnih izdelkov, presojanje, makropresoja, mikropresoja, dražljaji, pomembni za ocenjevanje

# 1 Introduction: The Dual Nature of Judgments

Current practice in the assessment of writing is grounded in an approach known as *judging*. The process of judging is understood here as involving raters not paying conscious attention to discrete elements of writing performances, and instead having them account for the overall impression that a writing performance leaves in the form of a rating (i.e., score or grade). The study at hand aims to demonstrate how an overall impression of any writing performance being evaluated in any assessment context, representing a (*macro*-)judgment of it made by a human rater, comprises, at its core, a complex and non-linear combination-count of individual (*micro*-)judgments. Each micro-judgment is itself triggered by individual *assessment-relevant* text features functioning as stimuli for the rater, highlighted by the expectations of a specific assessment context<sup>1</sup> and then being preconsciously recognized and mentally noted as such. Starting with existing rater cognition accounts of what assessment-relevant features of this kind may be, according to raters' own testimonies (Cai 2015; Chalhoub-Deville 1995; Cumming 1990; Cumming, Kantor, and Powers 2001; Eckes 2008; Milanovic, Saville, and Shuhong 1996; Mumford and Atay 2021; Sakyi 2000; and more), the study references one widely flagged and well-known type of individual assessment-relevant stimuli as an example, namely *errors*.

Because each instance of recognizing an error in a performance requires the rater to make a judgment relevant to both the operating language norm and to the specific assessment context as the pertinent sounding boards, the study argues that the overall macro-judgment of a writing performance is, at least partly, derivative from individually incurred micro-judgments triggered by raters recognizing errors as one set of influential stimuli. Other assessment-relevant stimuli (such as “assessment-positive” features, see Dobrić (2023)) can be assumed as being treated in a similar manner by raters.

Indeed, the earliest studies of judging that can be traced to the 1800s (Bejar 2012) saw the composite nature of judgments as central to how they are formed. Seminal early work came from Gustav Fechner, writing in 1863 about how an assessment of a work of art is conducted, based on the human ability to discriminate between shades of colour and similar phenomena. The ability is presumably activated by a particular combination of cues and stimuli stemming from the discernible features of the work of art being observed.

This initial research, together with his work that followed, such as Fechner (1897), is often credited with laying the groundwork for a number of different approaches to accounting for *aesthetic judgments* (e.g., Leder et al. 2004). The crux of this understanding is the proposition that observable, concrete features found in works of art form a basis for accounting for aesthetic evaluations (Bejar 2012). This direction of inquiry proved to be the catalyst for many other studies (Prall 1929; Peters 1942; Carlson 1981; and more), which together constitute norm-based or *analytical* (Sadler 1989, 129) accounts of the process of making judgments.

---

<sup>1</sup> The assessment context generally includes, on the one hand, the rater's own expectations, the relevant rating scale, and other externally prescribed evaluation-relevant criteria, and, on the other hand, the text being observed, all constituting the comparative framework in which a writing performance is rated.

The common core of analytical models of judging lies in understanding this process as an interaction between the rater as the information processor and sets of distinctive features serving as stimuli. One of the most significant models to emerge in this line of argumentation was the *lens* model (Brunswik 1952). Later providing the foundation for a major new perspective in psychology known as *Probabilistic Functionalism* (Postman and Tolman 1959), the lens model of judgments has been adapted to account for many research contexts in the humanities and social sciences (Engelhard, Wang, and Wind 2018).

Over half a century of applying and improving the model has resulted in the development of a general cues-based schema for describing human perception on the basis of the original model. In the centre of the process there is the latent variable that is being evaluated (Everitt 1984). An individual rater utilizes a set of related cues in a more or less unique fashion and provides a central response on the basis of the observation (Hammond 1955). In addition to the lens model itself and its different instantiations used to underpin analytic theories of judgments, further conceptual representations supporting the same idea can also be found. These include *decision trees* (Braun, Bejar, and Williamson 2006), the *normative rater cognition* model (Bejar 2011), and also the implicit models outlined by Suto and Greateorex (2006) and Crisp (2012).

Ultimately, the common thread that runs through the analytic accounts of how judgments are reached lies in envisaging the judging process as reducible to a formula comprising a number of formal criteria. This sweeping general claim and the different models that espoused it can be seen as empirically backed up by various studies. One such example of findings supporting this claim is the overlap discovered between rating scales, on the one hand, and the text features and qualities noted by raters as assessment-relevant, on the other (Milanovic, Saville, and Shuhong 1996). Additional examples include the many recorded cases where there is considerable overlap between raters agreeing on the quality of texts without formal recourse to rating scales (Morgan 1996).

When pointing out the downsides of this kind of modelling, researchers often state that the argumentation propping it up is limited to an excessive degree to the over-simplified linear weighing up of attributes (Bejar 2012, 3). Such a reduced process does not correspond to the complexity inherent in the rating task, which is closely bound up to an array of associated skills and abilities and varies subtly under different contexts of assessment. A modelling paradigm that is more composite in its approach thus seems to be called for. This complexity, with the problems it poses for the analytical models, was first addressed in the mid-20th century in studies revolving around *holistic* scoring (Coward 1952; Diederich, French, and Carlton 1961; Godshalk, Swineford, and Coffman 1966; Myers, McConville, and Coffman 1966; and more).

The apparently counterintuitive idea of a rater assigning a score to a writing performance holistically, based on a general impression, was not easy to accept as a substitute for the image of a deliberate and constructed analysis of a performance (Elliot 2005, 160). However, research showed that a less deliberate analysis reduces the cognitive effort and potentially increases rater agreement (Bejar 2012, 3). With this came the firm realization that much more is indeed involved in the judging process than an analytical model can

capture. This insight gave rise to the development of *configurational* theoretical accounts of judgments (Kaplan 1964, 187).

Configurational theories of judgments see the act of judging as entailing, to a far greater extent than recognized by any analytical model, a number of intuitive processes, which are, in turn, dependent on a multifaceted set of anything but linearly interplaying factors. This density can be explained by the fact that, in purely psychological terms, judging is a process of *bounded rationality* (Gigerenzer and Selten 2002; March 1994; Rubinstein 1997; Simon 1957). This means that human evaluation procedures are seriously impacted by our limited information processing capacity. This can be exacerbated by having to arrive at categorization decisions under conditions of uncertainty, such as is the case when there are not enough tangible criteria to guide the process. This is certainly true for assessment of writing and the rating scales used for this purpose.

However, in the context of researching the judging processes at work in the field of assessing writing performances, statements such as “I know a [2:1] when I see one” (Ecclestone 2001, 301) and the existence of significant levels of agreement over the quality of a text, even without any recourse to common rating scales or shared rater training (Morgan 1996; Wolf 1995), seem to support the view that judgments are not solely configurational in nature, either. Therefore, there is evidence that the process of judgment seems to be rule-governed at its core and activated by certain tangible stimuli resident in observed texts and expected by the operating assessment context and the operating raters, while there still being an unformalizable complexity fuelling the holistic dimension of judging.

Such reasoning leads to a consideration that there is a need for an understanding of judgments which conceptualizes the judging process invoked in rating a writing performance as seemingly intangible on the surface, but nevertheless structured in terms of deep processing, as a hybrid model. Such a model of judging would take equally from both the analytic and configurational approaches, and would follow the analytic conceptualization of judgments in as much as it would accept the existence of assessment-relevant stimuli in the observed responses (i.e., writing performances) as the formal basis of judgments. It would follow the configurational conceptualization by allowing for the interaction between raters and the assessment-relevant stimuli – as the formal basis of judgments – to be much more complex than can be expressed by simple counts or, most likely, by any kind of linear representation. The hybrid model would also allow for a complex interaction of other factors and raters, beyond the rating-relevant features alone,<sup>2</sup> thus acknowledging many potentially construct-irrelevant stimuli that may bring bias into the whole process. The plausibility of such hybridity in the process of judging is also vividly illustrated when judgments are revealed as being largely preconscious in nature.

---

<sup>2</sup> Such potentially construct-irrelevant influences can result from, for instance, the problems that rating scales bring in terms of vagueness of descriptors (Knoch 2011). Rater backgrounds may also lead to undesired score variability, together with the stylistic differences in the cognitive approaches of the raters (Scott and Bruce 1995). Other sources of bias can be rater strictness or potentially discriminatory behavior (Park, Chen, and Holtzman 2015). All of these can have been found to affect the decision-making and the rating of writing performances (Thunholm 2004).

## 2 Judgments as Preconscious Processes

Different studies within the educational field have tried to account experimentally for the cognitive nature of judging (Cooper and Odell 1977; Crisp 2010; Faigley 1989; Grainger, Purnell, and Zipf 2008; Tigelaar et al. 2005; and more). Most notably, there is the dual-processing theory of judgments proposed by Suto and Greateorex (2008). It focuses on registering the intensity of cognitive processing and follows from the awareness of the decision-making process which invariably takes place within the constraints of the limited human capacity to process information. Suto and Greateorex suggest that judgments can be classified as falling into one or other of two categories (2008, 215):

SYSTEM 1 judgments: judgments in this category are quick, associative, and often intuitive. The judging actions undertaken in this manner are preconscious and effortless. They require little to no recourse to rational thought and happen rapidly and in parallel. Examples of such judgments can be found in common statements such as the already mentioned “I know a [2:1] when I see one” (Ecclestone 2001, 301); and

SYSTEM 2 judgments: judgments in this latter category are slow, rule-governed, and raters are aware of employing the relevant rules. Such judgments are systematic, they require a lot of cognitive effort, and are, in principle, fully consciously controlled by the rater. However, it is important to note that for individual raters SYSTEM 2 actions tend to become SYSTEM 1 actions with increased experience. This is evident in the attested fact that the more experienced the raters are, the less likely they are to overtly follow prescribed rules (Wolf 1995).

The existence of two broad categories for the types of mental processes involved in the act of rating (judging), one more intuitive and the other more rational, is also indicated by a number of similar studies, such as Milanovic, Saville and Shuhong (1996). It suggests the presence of an initial *skimming* phase, when raters assign a provisional internal score, followed by an *in-depth* phase, in which raters come back to their own initial score and double-check. This latter phase basically includes comparing the original intuitive judgment based on first impressions with prescribed rating criteria and thus serves as an internal *monitoring* or an internal standardization procedure. An analogous division is also posited by Cooksey, Freebody and Wyatt-Smith (2007), Ecclestone (2001), Keren and Teigen (2004), Lumley (2005), and Shirazi (2012), among others. In short, the focus of this line of research is on clearly differentiating between the rational thought processes and the intuitive responses involved in judging (Brooks 2009).

The challenge facing researchers investigating the judging process is one of accounting for the process of making tacit knowledge explicit (Eraut 2000, 119). Linking judgments to *heuristics* is therefore also in line with this kind of research. As a process, heuristics reduce cognitive complexity and represent a means of facilitating a *mental shortcut* (Gilovich and Griffin 2002, 3). Heuristics function by reducing uncertainty via a reference to previously established “anchors” (Mussweiler and Englich 2005; Tversky and Kahneman 1992). External pressures from the volume of work, time constraints, and the like are often additional stress factors in evaluation. Therefore, it is easy to understand the tendency of raters to either deliberately or pre-cognitively create strategies which facilitate shortcuts in order to complete the judgment

process in the most resource-economical way possible (Gigerenzer and Goldstein 1996). The ultimate goal, motivated both from within, cognitively, and from without, as resource-based pressure, seems to be to minimize the intellectual effort and investment of resources but to still get the job done (Krosnick 1991; Quinlan, Higgins, and Wolff 2009; Simon 1957).

What can be concluded from this discussion is that arriving at judgments can be seen as a combination of rapid, preconscious processes (SYSTEM 1) and conscious rule-governed processes (SYSTEM 2). The central tendency, nevertheless, is for the structured explicit processes to become intuitive, an evolution governed by limited cognitive handling capacity and driven by the increasing experience and expertise of the raters. Less experienced raters tend to focus more consciously on specific quality-relevant features. With experience, however, they tend to approach texts in a more pre-cognitive manner and to ultimately adopt a more holistic approach (Furneaux and Rignall 2007).

Nevertheless, in the migration from conscious, more analytic processing (SYSTEM 2) to the more preconscious, holistic approach (configurational, SYSTEM 1), it seems that only the manner of processing is changed (from slow and calculated to fast and intuitive) rather than what actually drives the process. There remains something largely consistent underlying congruent judgments made by different raters. This seems to indicate that the psychological perception of judgments also confirms the existence of both analytic and a configurational side to the judging process, in congruence with the discussion offered above. A similar conceptualization of judging also appears to stem from the perception of judgments as complex comparison procedures.

### 3 Judgments as Comparisons

That the forming of judgments, with the intricate processing of information which this entails, is suitable to be imagined as involving a complex procedure of comparison (Laming 2004, 9) was evident even in the earliest models of judging, such as the previously referenced lens model (Brunswik 1955). Here, comparisons were being made between certain activated mental representations and certain salient features of the observed phenomenon. In the context of rating writing performances, judgments can, in this respect and as a starting point, be conceptualized as essentially *reading comprehension* procedures. The meaning of the text is actively constructed, based on the interaction between the readers' knowledge and the content supplied by the text (Johnson-Laird 1983, 402).

Cumming (1990, 33) similarly suggests the existence of two general clusters of rating attention when it comes to writing assessment, namely one concerned with reading and another, *interpretive* one. The actual decision-making occurs as a mode of retrieving relevant information from memory and comparing it to the meaning constructed on the basis of the observable textual input (Baker 2012, 226). In this respect, rating can be conceptualized as the process of comparing the knowledge structures built on the basis of the observed performance with the rater's own internally activated corresponding mental representations of the self-same performance (Crisp 2008), generated by the rater's own expectations and largely motivated by the expectations of the assessment context.

Even when the description of the rating process is not confined to accounting for only a few general steps, but instead is more finely grained, it can still be conceptualized in terms of making comparisons. For example, Bejar describes the scoring procedure as a sequence of several phases (2012, 5):

- (1) the rater first reads the writing performance and produces a mental representation of it, a process which takes the form of a comparison between the context-conditioned *hypothetical “ideal”* instantiation of the observed text and the observed text itself;
- (2) the rater then weights the previously formed mental representation in relation to a *mental scoring rubric*. The mental scoring rubric represents the rater expectations in more explicit form, including potentially internalized rating criteria; and
- (3) the rater assigns the performance to a score category, possibly after a self-monitoring phase which can include a more conscious reference to the externally prescribed criteria, referenced by experienced raters most commonly for the purposes of uniformity of expression rather than for modifying the impression already formed.

Myford and Wolfe view judging as a process of information acquisition and processing, punctuated by a number of rater functions (2003, 37). These include: (1) the rater functioning as an information processor; (2) the rater retrieving information from memory storage; and, ultimately, (3) the rater combining, weighting, and integrating the information to draw rating inferences. Wiliam also lists several comparands, including *criteria*, *constructs*, *self*, and *norms* (1992, 17). DeRemer conceptualizes the rating process as a *problem-solving exercise* (1998, 27), including (1) the *simple recognition task* elaboration (i.e., general impression scoring); (2) the *rubric searching* elaboration (i.e., rubric-based evaluation); and (3) the *complex recognition* elaboration (i.e., text-based evaluation). Finally, synthesizing a myriad of accounts, Gauthier, St-Onge and Tavares list seven general mechanisms involved in making judgments, which can be grouped into the three phases of (1) observation (reading); (2) processing; and (3) integration (2016, 511). These mechanisms comprise:

- (1) generating automatic impressions;
- (2) formulating high-level inferences;
- (3) focusing on different dimensions of the competencies of interest;
- (4) categorizing through well-developed schemata, based on personal concepts of the given competence, comparison with various examples, and also task and context specificity;
- (5) weighting and synthesizing information in a number of different ways;
- (6) production of mature judgements; and
- (7) translating narrative judgments into scales.

On the part of the rater, this comparison process can be affected by several additional factors. These include cognitive constraints (Simon 1957, 227), prior knowledge and experience (Newell, Lagnado, and Shanks 2007), and possibly a number of potentially *construct-irrelevant* influences (i.e., *rater effects*, see Eckes (2023)). The cognitive constraints, as already noted, are the limiting factors in human cognition that are responsible for ultimately making

judgments, to a large extent, the result of preconscious intuitive processes. Prior knowledge and experience are embodied, as a combination with externally prescribed criteria, into the *fuzzy* hypothetical “ideal” form a particular performance should take in a particular context following a particular prompt.

This firstly corresponds to the mental representation that is activated rater-internally, existing *in potentia* in the rater’s mind and instantiated only as an *authoritative reconstruction* of the observed text, embodying the rater’s expectations. This rater-internal point of comparison can be further influenced by various external and co-opted rating criteria that a rater has been asked to use, and through experience, has potentially internalized and learned to apply (Suto and Greatorex 2008, 215). The fact is that, in the course of the migration from self-aware to preconscious judging, the externally prescribed rating criteria can undergo change and become adapted to fit the already existing rater expectations (Brooks 2009, 12). Finally, rater expectations, commonly incorporating external rating criteria, can be further affected by various factors such as cultural background, rater strictness, and different sources of bias (Hay and Macdonald 2008, 159).

That making comparisons is central to the nature of judgments is also revealed by studies looking into the *weighting patterns* that raters exhibit. Cumming (1990, 47) shows that among 28 identified types of rater behaviour, raters, overall, focus on comparisons based on *self-monitoring* on the one hand and *rhetorical*, *ideational*, and *language* content on the other. Chalhoub-Deville, in her discussion of deriving assessment scales for oral performances, similarly identified *grammar-pronunciation*, *creativity* in presenting information, and *amount of detail* provided as relevant comparands (1995, 48). Cai (2015) found raters putting emphasis on either form or content, or on achieving a balance between the two. Eckes (2008) categorizes raters according to the weight they attach to *syntax*, *correctness*, *fluency*, *non-fluency*, and *non-argumentation*. Baker (2012), in turn, classifies raters as regards their emphasis style being *rational*, *intuitive*, *dependent*, *avoidant*, and *spontaneous*.

In summary, it can be concluded that conceptualizing judgments as comparisons also seems to confirm the existence of concrete features in texts which serve as stimuli for rater cognition and, ultimately, for categorization decisions. Central to the process of comparison are, in fact, discrete aspect of texts which are recognized by raters as being in line or not in line with their (assessment) expectations relevant for a particular educational/assessment context and relative to a particular performance being observed. It is not surprising, therefore, that much research effort has been invested into describing what these discrete aspects of texts may comprise.

## **4 Assessment-Relevant Stimuli Responsible for Triggering Micro-Judgments – Errors as Prototypical Cues**

Common to the presented conceptualizations of the dual nature of judgements (i.e., configurational vs. analytical accounts, the dual-processing view on the underlying cognition, and the conception of judgments as comparisons) is their direct or indirect focus on trying to pinpoint concrete types of assessment-relevant text cues and stimuli by which the raters may be influenced when assessing writing. This has, in fact, always been one of the central questions in rater cognition research.

For example, Milanovic, Saville and Shuhong identified 11 essay elements constituting rater focus, with different weighting attached to each by the informants, whereby these included *grammar, communicative effectiveness, content, length, legibility, spelling, structure, task realization, tone, punctuation, and vocabulary* (1996, 106). Similarly, the study conducted by Sakyi identified four concrete focus points, including *errors in the text, essay topic and presentation of ideas, raters' personal reactions to the text, and prescribed scoring criteria* (2000, 140). Cumming, Kantor and Powers report clustering of rater focus regarding three broad categories of text features, including *rhetorical organization, expression of ideas, and accuracy and fluency* (2001, 3).

The general conclusion from these and other similar studies, stemming mainly from interviews and think aloud protocols, seems to be that raters essentially focus on three major qualities of a text, which can be loosely termed correctness, content representation, and relation to the task. This has indeed been recognized by the rich field of *CAF* research (Skehan 2009), with studies stipulating *complexity* (as expectations of mature language performance, see DeKeyser (2008)), *accuracy* (as expectations of norm-adequacy, as in Xie (2019)), *fluency* (as expectations of smoothness or ease of speech or writing, see Hilton (2008)), and *functional adequacy* (as expectations of task achievement, in Kuiken and Vedder (2017)), as the main cornerstones of language competence.

What such studies generally do not focus on, however, is that the crucial aspect of assigning assessment-relevance to any text feature involves making the distinction of whether it has been performed successfully and in alignment with the operating assessment expectations or not (Dobrić 2024). It is not enough just to catalogue the different types of features that can be assessment-relevant. The manner of identifying the direction and magnitude of their assessment-relevance when observed in action, appearing in a concrete performance, is also necessary to consider. For instance, every time a text feature is recognized by a rater as not performed as expected, regardless of whether this is in relation to accuracy, complexity, fluency, or functional adequacy (thus being highlighted as *inaccurate* and/or *inappropriate*), a certain degree of *assessment-negative* value is usually recognized for it, too. In other words, we have a case of an error being registered (i.e., micro-judged), triggering on the instance of an assessment-relevant stimulus (usually in a negative direction) on the side of the rater.

Errors have been referenced for generations as text cues relevant to writing assessment, typically found underlined using the dreaded red pen and often implied as constituting much of the basis of the grade received on a writing task. An error, in this sense, is understood as any text feature which does not conform to rater expectations relevant for one assessment context. This includes both the more traditional understanding of errors as related to correctness and *non-norm adequacy* (Richards 1974) and, far beyond any formalizable norm, those related to *inappropriateness*, relative to performance expected in fulfilment of a specific assessment task in a specific setting.

Despite this, “error counting” approaches to the rating of writing performance are generally viewed very negatively by the applied linguistics field, and it is not difficult to see why. Namely, an emphasis solely on errors in the assessment of productive language skills is likely to be counter-productive, misrepresentative, and demotivational. A “good” text entails significantly

more than it only being “error-free”. Still, reference to errors as important assessment-relevant cues and stimuli persists in educational practice, for several reasons (Dobrić and Sigott 2023).

Firstly, an attention to errors provides raters with an intuitively transparent set of assessment criteria, generally offering much more structure and clarity than the often rather vague rating scale descriptors. Secondly, the usefulness of errors also stems from them being conceptualized as indicators of problems in the learning process, (Dobrić 2015; Dulay, Burt, and Krashen 1982; Havranek 2002; James 1998; Lengo 1995; Lennon 1991; Swan and Smith 2001). This is what makes errors so indispensable for the provision of *constructive feedback* (Sigott et al. 2019).<sup>4</sup>

For all these reasons, it is relatively safe to claim that errors represent a prototypical case of a type of assessment-relevant set of text features generally influential when making judgments of writing performance (Dobrić and Sigott 2014). They are one of the key features that facilitate the comparison between the rater’s mental representation of an observed text and the observed text itself. Therefore, the manner in which errors, for whose recognition a significant amount of comparison (i.e., judging) is required in the first place, go on to help facilitate this process of complex comparison is key to understanding the nature of their contribution to forming judgments (and, for that matter, of the contribution of any type of conceptually comparable assessment-relevant text cues or stimuli one may propose beyond errors).

## 5 Micro-Judgments into Macro-Judgments

The manner in which errors, as a well-established and widely recognized type of assessment-relevant text features, can be understood as contributing to judgments has been illustrated in the manner of how judgments largely function as fast, preconscious comparisons. Errors, as one participant in this complex comparison procedure, tend to indicate “undesirable” divergences between, on the one hand, the written performance that is actually observed and, on the other hand, the written performance expected by the rater and the surrounding assessment context. This weighting of similarities and differences between rater-internal expectations and rater-external stimuli,<sup>5</sup> or in other words, the comparison between the hypothetical ideal version of a text (for a set context) and the self-same text as actually performed, results in a “vector of similarities” being constructed using these two broad comparands (Bejar 2012, 5).

This vector is a result of a reconciliation taking place between the rater’s overall impression of a text (as reflected in their expectations), the specific yet often hard to pinpoint, assessment-relevant text features (often even explicitly recorded for feedback purposes), and, finally, any given (external or internalized) rating scale criteria (Zhang 2016). Ultimately, the vector of similarities ends up representing the overall judgment of a text and directs the predisposition,

---

<sup>4</sup> In truth, the appeal of including errors in the assessment process may go even further, perhaps answering to the human instinct to associate any kind of evaluation first and foremost with pointing out faults. This is observable in many human-based evaluation practices, such as academic project reviews or professional performance evaluations.

<sup>5</sup> Explicit attention should once more at this point be drawn both to the fact that this process of comparison can be additionally influenced by various sources of potential bias, and that this process is anything but as linear or straightforward as may be inadvertently suggested by the narrative.

in the mind of the rater, to classify a performance as belonging to a particular quality-related category by assigning it a rating.

The novel insight important to highlight at this point – stemming from how errors have been accounted for as one set of text features raters themselves often report as relevant when assessing, and from how judging has been described as largely a preconscious procedure – is that the composition of the vector of similarity depends, by default and to a large extent, on individual instances of comparison on the text feature level, building up its geometry as coordinates. Namely, as Figure 1 shows, individual instances of comparison between text features (as assessment-relevant stimuli) expected and text features observed cumulatively (and, again, non-linearly) build up into an overall comparison being formed between a writing performance expected and writing performance observed. They thus construct the said vector of similarity which is ultimately translated into a rating.

The effects of any assessment-relevant text cue or stimulus are relative to the specific assessment expectations operating in a particular context (including here also the relevant language norm, see Gilquin (2022)). This makes them essentially *latent* phenomena, with each individual instance of comparison on a text-feature level, key to recognizing any error, requiring a judgment to take place on the side of the rater. A micro-judgment, as termed and understood here, thus involves any one individual text feature in a writing performance being (mostly preconsciously) recognized as assessment-relevant (alongside its negative or positive direction and magnitude of effect) by a human rater. It is flagged thus for the rater's own benefit and for the purpose of quality categorization by virtue of the direction and extent of its accuracy and/or acceptability (or lack thereof) in a specific context.

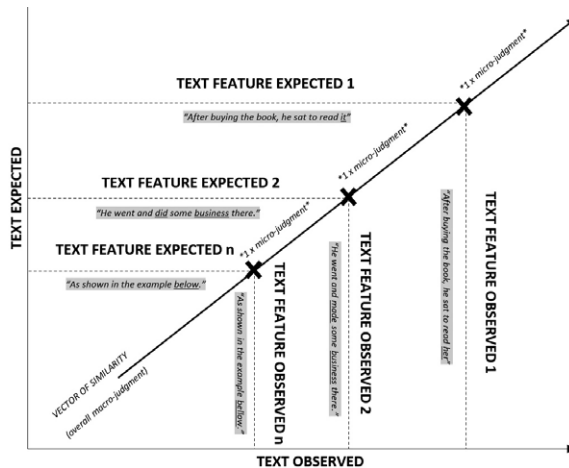


FIGURE 1. Simplified illustration of the role of errors as triggers for micro-judgments.

What makes errors ideal exemplification candidates in this case is that, as indicated earlier, the process of recognizing one text feature as an error (similar, in fact, to recognizing any text feature as assessment-relevant) is the same in structure as described previously in relation to forming an overall judgment of a writing performance, as illustrated in Figure 1. The process

requires the rater to judge the conformity of any given text features in terms of its correctness and acceptability, and it entails all of the complexity discussed in the previous sections, only on a smaller scale, involving one assessment-relevant text feature at a time.

Correctness, as one coordinate of rater expectations relevant for identifying errors, references the more codifiable and, hence, the more firmly codified aspects of language performance (i.e., language norm and standards). It is usually incorporated into the assessment expectations of all educational contexts involving language. Inappropriateness, on the other hand, does not relate to any firmly codified (or even codifiable) aspects of language performance, but rather to often gradable judgments made as to the degree to which particular text features conform to particular operating (general and specific, text-internal and text-external) assessment expectations (Dobrić et al. 2021).

For instance, if we take the “After buying the book, he sat to read her” example sentence used in Figure 1 to illustrate the ‘TEXT FEATURE OBSERVED’ step of a micro-judgment being triggered, the manner in which the text feature her is detected in it as an assessment-relevant cue of negative direction (i.e., as an error) has to do with it not fitting to what a rater would expect when guided by expectation of English language competence relevant to anything but lowest levels of proficiency. The expectation, pointed out by the “After buying the book, he sat to read it” is one of suitable anaphoric reference made by the underlined pronoun and is motivated by both appropriateness and correctness as reference points. The micro-judgment that takes place is mirrored in the rater engaging in the comparison between the expected and perceived and deciding on how to interpret any discrepancy. In this case, this would take the direction of her being judged to be an error.

This and the other examples provided in Figure 1 solidify the claim that an overall judgment of a writing performance, which has been, in the light of this argumentation, termed a macro-judgment, is comprised of accumulated and in some way combined individual micro-judgments, when it comes to errors as stimuli at the very least. This accumulation process is, as indicated above, complex and involves an as of yet unmapped scheme of non-linear “counting and combining” of the micro-judgments into a macro-judgment. Nevertheless, the research indicates that the initial result of this process of macro-judging is also achieved preconsciously and generally takes the form of a fairly rapid decision on the categorization of the text (taking place as a SYSTEM 1 judgment, as in Suto and Grotorex (2008)). This is despite the fact that the final step in the process is one potentially involving some conscious deliberating, as a *self-monitoring* stage, during which external criteria (i.e., the rating scale) may be additionally consulted, in a more conscious and self-aware manner, prior to awarding the rating. In addition, as also already underlined several times, different biasing and presumed not construct-relevant factors may also influence the construction and direction of the relevant vector of similarity.

Hence, as Figure 1 illustrates in a necessarily simplified manner, in order to recognize a writing performance being assessed as conforming or not conforming (and to what degree) to the expectations operating in one assessment context, raters first judge each and every text feature in it (individually and in combination) to the same extent of potential conformity. This, as many studies have indicated, usually happens preconsciously (especially with experienced

raters), leaving the manner in which these lower-level micro-judgments are combined into an overall macro-judgment still largely opaque. When it comes to errors as a prominent and traditionally influential type of assessment-relevant features, the micro-judgments occur at the moment in which any text feature expected to appear in the text observed is not found performed in the anticipated (i.e., correct and/or appropriate) fashion or is perhaps missing.

Other assessment-relevant factors identified in the different studies listed before, besides errors, presuppose additional types of assessment-relevant text features being postulated, presumably constituting a group that can be interpreted as assessment-positive, i.e., as diametrically opposite to errors which are assessment-negative in effects (Dobrić 2023). These can be expected – themselves by default also representing latent phenomena – to also assert influence via the micro-judgments they trigger, only in the direction contrary to that of errors (i.e., with a positive value, steering raters towards awarding a higher rather than a lower rating).

## 6 Errors and Judgments – A Case Study

Empirical backing for the assertion of the prototypical role of errors in forming micro-judgments which consequently inform (together with other kinds of assessment-relevant text features) the macro-judgments of writing performances offered above can be found in a recently published study by Dobrić (2024). The research, focusing on the effects of errors on rating writing performance, demonstrates how errors explained roughly 50% of overall score variability in a high-stakes EFL context. This was found despite the fact that the rating criteria (i.e., the rating scale) employed in the given context actually directed the raters towards a focus on assessment-positive aspects of texts (rather than the assessment-negative qualities of texts that are represented as errors). Table 1 briefly recaps the reported results.

TABLE 1. Pseudo  $R^2$  coefficients (Nagelkerke 1991)<sup>6</sup> reported in the model, representing the overall effects of errors on the different rating dimensions of the employed analytic rating scale (Dobrić 2024).

Ratings	Nagelkerke Pseudo R-Square	Variability of Ratings Accounted For
Vocabulary	0.358	36%
Grammar	0.617	62%
Textual Competence	0.569	57%
Aggregated ratings	0.512	51%
average	0.514	51.5%

What the study demonstrates, in terms of how judgments were formed in the case of the high-stakes writing exam observed in the study, is that a concrete set of distinctive text features (which can be argued as assessment-relevant) could be empirically brought into correlation with the independently formed rater judgments (expressed as grades). In other words, even

<sup>6</sup> Pseudo  $R^2$  coefficient is a supplementary measurement used to demonstrate the compound effect of independent (i.e., predictor) variables in explaining the variance of the dependent (i.e., predicted) variable on the whole, emulating the more familiar  $R^2$  coefficient from linear regression (Veall and Zimmermann 1996).

though the raters who originally processed the writing performances analysed in Dobrić (2024) received both overt and covert instruction (i.e., by means of rater training and via the rating scale employed for that purpose) in how to engage in judgments (i.e., employ a more configurational rather than analytical approach), their decisions could still be quantitatively (i.e., by means of logistic regression) linked, in a causal manner, to individual features of the rated texts (i.e., in this case to errors).

As already argued above, each of the individual features of a text comprising an instance of an error demanded a judgment to be made as to it actually constituting an error in the first place. Considering that the analysed writing exam took place at a university department of English and involved C1–C2 expectations of L2 English proficiency, there were many errors which demanded decisions being made in terms of their appropriateness rather than of their correctness (Dobrić and Sigott 2023). What this lucidly evokes is the image of micro-judgments, triggered by errors, cumulatively informing macro-judgments, expressed as grades, as illustrated in the section above. In the case of Dobrić (2024), this was attested to be at the rate of between 36% and 62% of the variability of the macro-judgments (as indicated by the pseudo  $R^2$  values provided in Table 1) likely informed by error-based micro-judgments.

What is more, the fact that it was around 50% of the variability of the grades observed in the examined high-stakes university writing exam covered by Dobrić (2024) indicates that something else is responsible for the unexplained portion of the variability of the self-same grades. It is assumed but not yet empirically demonstrated that the proposed assessment-positive text features (Dobrić 2023) may play a significant role in this respect, possibly responsible for macro-judgments in a similar way as errors when triggering micro-judgments, only of a different (i.e., opposite) assessment value and direction.

## **7 Consequences of the Findings for Understanding Writing Assessment**

Even though the exact manner in which micro-judgments (and possibly other factors) interact and come together to form overall macro-judgments (ultimately filtered through rating scales into ratings) is still largely unaccounted for (and very difficult to formalize, being of both analytical and configurational nature), the realization that distinctive cues and stimuli found in assessed writing performances also require the activation of a judgment process is a noteworthy one. It is not only relevant in terms of contributing to our understanding of rater cognition but is also important, in a practical sense, in terms of rater training and the construction of rating scales.

Starting with rater cognition, the realization that an overall judgment of a writing performance is comprised of many, mostly preconsciously made, smaller-scale judgments is significant because it reveals a whole new level at which subjectivity may be affecting measurements of writing competence, thus highlighting the need for accounting for this in some manner (similarly to how rater effects are accounted for at the level of macro-judgments and scoring, see Eckes (2023)). Namely, just the example of errors as one possible type of assessment-relevant text features likely responsible for triggering micro-judgments indicates the issue of *gradience*

intrinsic to recognizing the majority of error types and of their potential assessment-relevance and, by extension, the amount of potential subjectivity involved (as often demonstrated in by studies looking into *inter-annotator agreement*, see Sigott, Cesnik, and Dobrić (2016)).

The issues of gradience and subjectivity are only aggravated when other types of assessment-relevant text features that may be postulated are considered, such as the hypothetical assessment-positive ones (Dobrić 2023). Highlighting concretely what makes one writing performance “good” (i.e., acceptable or even extraordinarily acceptable) in terms of concrete text features and qualities related to one assessment context is, in contrast to errors – which do at some level also respond to the much more formalized correctness as a reference point – almost exclusively an act resident in judging. This is because determining assessment-positiveness involves solely deciding on the levels of adequacy (or excellence) of individual aspects of a text for the set context and not of correctness at all, possibly adding even more bias and potential for measurement error to the process.

This is why it is crucial to acknowledge the existence of an additional underlying level of potential subjectivity involved in rating writing (Dobrić 2018) and cast it into methodological solutions which would allow us to account for it when it comes to the rating, rater effects, and potential measurement errors it can produce. Similarly, it is also important to cast the outlined findings into rating practice by making the raters and rating criteria more sensitive to them (Cushing Weigle 1994; Dobrić et al. 2021). In other words, while methods for detecting and measuring rater effects in performance assessments have for a long time been a fixture of psychometric research on rating quality in the assessment of writing (Eckes 2015),<sup>7</sup> the diverse methodology stops at the level of the text as the measured phenomenon and the macro-judgment as the relevant process. The discussion so far has, however, shown that there is a need to account for measurement errors being produced below the text as a whole, in both testing research and practice. One step towards doing so is to utilize methods designed for describing annotator behaviour in learner corpus coding (Dobrić 2022). The myriad of methods devised for measuring intra- and inter-annotator agreement, usually revolving in some way around Kappa coefficients (McHugh 2012), could prove very useful for revealing the rater effects potentially present at text feature levels. This may include, similarly to those found at the text level, the raters when micro-judging individual textual features showing evidence of inaccuracy, illusory halo, severity/leniency, central tendency/extremity, and so on.

Then, in rater training, more emphasis needs to be placed on having raters become more aware of concrete text features and textual qualities that may be influencing their impression-making when it comes to writing performance (Cushing Weigle 1994). This can take the form of different exercises focused on identifying assessment-relevant cues and stimuli in example texts and then accounting for why they may be signifying relevance in the first place. Furthermore, emphasis should be placed on the different reference points necessary for the raters to employ in order to recognize a text feature or quality as assessment-relevant, such as the norm, the rating criteria, the authoritative reconstruction, and similar.

<sup>7</sup> Revolving largely around many-facet Rasch measurements, they are designed to detect and somehow account for measurement error human raters bring to the table when assessing writing. These include inaccuracy, illusory halo, severity and/or leniency, central tendency and/or extremity, and more (Wolfe and Song 2016).

Open for further research, and much more controversial than the proposition of the existence of micro-judgments, is the proposition that the manner in which they interact to comprise the macro-judgments (alongside bias as potential additional factors) can be analytically presented. This is precisely where the most prominent opportunities for study are to be found, especially in terms of natural language processing (NLP), artificial intelligence (AI), and algorithmic representations of rater cognition. Dismantling the complex cognitive process in which the vector of similarity between the expected writing performance and the observed writing performance is actually created in the rater's mind and doing so in a way which would not be solely reliant on raters' own testimonies (fraught with problems of *reactivity* and *veridicality* of informant responses, see van Someren, Barnard, and Sandberg (1994)), would be the next natural and necessary step to take in our understanding of what raters do when they assess writing and in our attempt to formalize it.

In reference to these open questions, additional studies would also be welcome in the direction of investigating further the different types of assessment-relevant text features other than errors that can be demonstrated as also responsible for triggering macro-judgments. Moreover, studies into methods which would allow for a better understanding of rater effects and of measurement errors potentially occurring at the text feature and micro-judgment levels, would also prove a promising research direction to take. Finally, studies which focused on cataloguing assessment-relevant features commonly appearing in specific contexts and using the findings for improving rating scales and rater criteria in the self-same settings, would also be of major academic and practical significance.

## References

- Baker, Beverly Anne. 2012. "Individual differences in rater decision-making style: An exploratory mixed-methods study." *Language Assessment Quarterly* 9 (3): 225–48. <https://doi.org/10.1080/15434303.2011.637262>.
- Bejar, Isaac. 2012. "Rater cognition: Implications for validity." *Educational Measurement Issues and Practice* 31 (3): 2–9. <https://doi.org/10.1111/j.1745-3992.2012.00238.x>.
- . 2011. "A validity-based approach to quality control and assurance of automated scoring." *Assessment in Education: Principles, Policy and Practice* 18 (3): 319–41. <https://doi.org/10.1080/0969594X.2011.555329>.
- Braun, Henry, Isaac Bejar, and David Williamson. 2006. "Rule-based methods for automated scoring: Application in a licensing context." In *Automated Scoring of Complex Tasks in Computer-Based Testing*, edited by David Williamson, Robert Mislevy, and Isaac Bejar, 83–123. Mahwar: Lawrence Erlbaum.
- Brooks, Val. 2009. "Marking as judgment." *Research Papers in Education* 27 (1): 63–80. <https://doi.org/10.1080/02671520903331008>.
- Brunswik, Egon. 1952. *The Conceptual Framework of Psychology*. Chicago: University of Chicago Press.
- . 1955. "Representative design and probabilistic theory in a functional psychology." *Psychological Review* 62: 193–217.
- Cai, Hongwen. 2015. "Weight-based classification of raters and rater cognition in an EFL speaking test." *Language Assessment Quarterly* 12 (3): 262–82. <https://doi.org/10.1080/15434303.2015.1053134>.
- Carlson, Allen. 1981. "Nature, aesthetic judgment, and objectivity." *The Journal of Aesthetics and Art Criticism* 40 (1): 15–27. <https://doi.org/10.2307/430349>.
- Chalhoub-Deville, Micheline. 1995. "Deriving oral assessment scales across different tests and rater groups." *Language Testing* 12 (1): 16–33. <https://doi.org/10.1177/026553229501200102>.
- Cooksey, Ray, Peter Freebody, and Claire Wyatt-Smith. (2007). "Assessment as judgement-in-context: Analysing how teachers evaluate students' writing." *Educational Research and Evaluation* 13 (5): 401–34. <https://doi.org/10.1080/13803610701728311>.

- Cooper, Charles, and Lee Odell, eds. 1977. *Evaluating Writing: Describing, Measuring, Judging*. Urbana: National Council of Teachers of English.
- Coward, Ann. 1952. "A comparison of two methods of grading English compositions." *Journal of Educational Research* 46 (2): 81–94.
- Crisp, Victoria. 2010. "Judging the grade: Exploring the judgement processes involved in examination grading decisions." *Evaluation & Research in Education* 23 (1): 19–35. <https://doi.org/10.1080/109500790903572925>.
- . 2012. "An investigation of rater cognition in the assessment of projects." *Educational Measurements: Issues and Practice* 31 (3): 10–20. <https://doi.org/10.1111/j.1745-3992.2012.00239.x>.
- . 2008. "The validity of using verbal protocol analysis to investigate the processes involved in examination marking." *Research in Education* 79 (1): 1–12. <https://doi.org/10.7227/RIE.79.1>.
- Cumming, Alistair. 1990. "Expertise in evaluating second language compositions." *Language Testing* 7: 31–51. <https://doi.org/10.1177/026553229000700104>.
- Cumming, Alistair, Robert Kantor, and Donald Powers. 2001. *Scoring TOEFL Essays and TOEFL 2000 Prototype Writing Tasks: An Investigation into Raters' Decision Making and Development of a Preliminary Analytic Framework (TOEFL Monograph Series, MS-22)*. Princeton: Educational Testing Service.
- Cushing Weigle, Sara. 1994. "Effects of training on raters of ESL compositions." *Language Testing* 11 (2): 197–223. <https://doi.org/10.1177/026553229401100206>.
- DeKeyser, Robert. 2008. "The complexities of defining complexity." Paper presented at AAAL 2008, April 1st, 2008. Washington D.C.
- DeRemer, Mary. 1998. "Writing assessment: Raters' elaboration of the rating task." *Assessing Writing* 5 (1): 7–29. [https://doi.org/10.1016/S1075-2935\(99\)80003-8](https://doi.org/10.1016/S1075-2935(99)80003-8).
- Diederich, Paul, John French, and Sydel Carlton. 1961. "Factors in judgments of writing ability." *ETS Research Bulletin* 61: 2.
- Dobrić, Nikola. 2024. "Effects of errors on ratings of writing performances in rating contexts in which focus on them is overtly discouraged – Evidence from a high-stakes exam." *Assessing Writing* 59: 100806. <https://doi.org/10.1016/j.asw.2023.100806>.
- . 2023. "Towards a taxonomy of text features highly indicative of context-appropriate L2 writing competence." *System* 119. <https://doi.org/10.1016/j.system.2023.103155>.
- . 2022. "Identifying errors in a learner corpus – the two stages of error location vs. error description and consequences for measuring and reporting inter-annotator agreement." *Applied Corpus Linguistics* 3 (1). <https://doi.org/10.1016/j.acorp.2022.100039>.
- . 2018. "Reliability, validity, and writing assessment: A timeline." *ELOPE: English Language Overseas Perspectives and Enquiries* 15 (2): 9–24. <https://doi.org/10.4312/elope.15.2.9-24>.
- . 2015. "Quality measurements of error annotation – ensuring validity through reliability." *The European English Messenger* 24: 36–42.
- Dobrić, Nikola, and Günther Sigott. 2023. "Use of error profiles in applied linguistics – Empowering language instruction by cataloguing rating-negative performance at the English Department, University of Klagenfurt, Austria." In *Power in Language, Culture, Literature and Education: The Perspectives of English Studies*, edited by Marta Degani and Werner Delanoy, 97–117. Tübingen: Günther Narr.
- . 2014. "Towards an error taxonomy for student writing." *Zeitschrift für Interkulturellen Fremdsprachenunterricht* 19 (2): 111–18.
- Dobrić, Nikola, Günther Sigott, Gasper Ilc, Vesna Lazović, Hermann Cesnik, and Andrej Stopar. 2021. "Errors as indicators of writing task difficulty at the Slovene general Matura in English." *International Journal of Applied Linguistics* 31: 475–91. <https://doi.org/10.1111/ijal.12345>.
- Dulay, Heidi, Marina Burt, and Stephen Krashen. 1982. *Language Two*. Oxford: Oxford University Press.
- Ecclestone, Kathryn. 2001. "I know a 2:1 when I see it: Understanding criteria for degree classifications in franchised university programmes." *Journal of Further and Higher Education* 25 (3): 301–13. <https://doi.org/10.1080/03098770126527>.
- Eckes, Thomas. 2023. "Detecting and measuring rater effects in performance assessments: Advances in many-facet Rasch modeling." In *Festschrift in Honour of Günther Sigott: Advanced Methods in Language Testing*, edited by Nikola Dobrić, Hermann Cesnik, and Claudia Harsch, 195–223. Frankfurt: Peter Lang.

- . 2015. *Introduction to Many-Facet Rasch Measurement: Analyzing and Evaluating Rater-Mediated Assessments*. 2nd ed. Frankfurt: Peter Lang.
- . 2008. "Rater types in writing performance assessments: A classification approach to rater variability." *Language Testing* 25 (2): 155–85. <https://doi.org/10.1177/0265532207086780>.
- Elliot, Jane. 2005. *Using Narrative in Social Research. Qualitative and Quantitative Approaches*. London: Sage Publications.
- Engelhard, George, Jue Wang, and Stefanie Wind. 2018. "A tale of two models: Psychometric and cognitive perspectives on rater-mediated assessments using accuracy ratings." *Psychological Test and Assessment Modelling* 60 (1): 33–52.
- Eraut, Michael. 2000. "Non-formal learning and tacit knowledge in professional work." *British Journal of Educational Psychology* 70: 113–136. <https://doi.org/10.1348/000709900158001>.
- Everitt, Brian. 1984. *An Introduction to Latent Variable Models*. Dordrecht: Springer Netherlands.
- Faigley, Lester. 1989. "Judging writing, judging selves." *College Composition and Communication* 40 (4): 395–412.
- Fechner, Gustav. 1897. *Kollektivmasslehre*. Leipzig: Engelmann.
- . 1863. *Die drei Motive und Gründe des Glaubens*. Leipzig: Breitkopf und Härtel.
- Furneaux, Clare, and Mark Rignall. 2007. "The effect of standardization-training on rater judgements for the IELTS Writing Module." In *IELTS Collected Papers: Research in Speaking and Writing Assessment SiLT 19*, edited by Lynda Taylor and Peter Falvey, 422–45. Cambridge: Cambridge University Press.
- Gauthier, Geneviève, Christina St-Onge, and Walter Tavares. 2016. "Rater cognition: Review and integration of research findings." *Medical Education* 50 (5): 511–22. <https://doi.org/10.1111/medu.12973>.
- Gigerenzer, Gerd, and Reinhard Selten. 2002. *Bounded Rationality. The Adaptive Toolbox*. Cambridge: The MIT Press.
- Gigerenzer, Gerd, and Daniel Goldstein. 1996. "Reasoning the fast and frugal way: Models of bounded rationality." *Psychological Review* 103 (4): 650–69. <https://doi.org/10.1037/0033-295X.103.4.650>.
- Gilovich, Thomas, and Dale Griffin. 2002. "Introduction – heuristics and biases: Then and now." In *Heuristics and Biases. The Psychology of Intuitive Judgement*, edited by Thomas Gilovich, Dale Griffin and Daniel Kahnemann, 1–19. Cambridge: Cambridge University Press.
- Gilquin, Gaëtanelle. 2022. "One norm to rule them all? Corpus-derived norms in learner corpus research and foreign language teaching." *Language Teaching* 55 (1): 87–99. <https://doi.org/10.1017/S0261444821000094>.
- Godshalk, Fred, Frances Swineford, and William Coffman. 1966. *The Measurement of Writing Ability*. New York: College Entrance Examination Board.
- Grainger, Peter, Ken Purnell, and Reyna Zipf. 2008. "Judging quality through substantive conversations between markers." *Assessment & Evaluation in Higher Education* 33 (2): 133–42. <https://doi.org/10.1080/02602930601125681>.
- Hammond, Kenneth. 1995. "Probabilistic functioning and the clinical method." *Psychological Review* 62 (4): 255–62.
- Havranek, Gertraud. 2002. *Die Rolle der Korrektur beim Fremdsprachenlernen*. Frankfurt/Main: Peter Lang.
- Hay, Peter, and Doune Macdonald. 2008. "(Mis)appropriations of criteria and standards-referenced assessment in a performance-based subject." *Assessment and Evaluation in Higher Education: Principles, Policy and Practice* 15 (2): 153–68. <https://doi.org/10.1080/09695940802164184>.
- Hilton, Heather. 2008. "The link between vocabulary knowledge and spoken L2 fluency." *Language Learning Journal* 36 (2): 153–66. <https://doi.org/10.1080/09571730802389983>.
- James, Carl. 1998. *Errors in Language Learning and Use: Exploring Error Analysis*. London: Longman.
- Johnson–Laird, Philipp. 1983. *Mental Models: Towards a Cognitive Science of Language, Inference, and Consciousness*. Cambridge: Cambridge University Press.
- Kaplan, Abraham. 1964. *The Conduct of Inquiry: Methodology for Behavioural Science*. San Francisco: Chandler Publishing Company.
- Keren, Gideon, and Karl Teigen. 2004. "Yet another look at the heuristics and biases approach." In *Blackwell Handbook of Judgment and Decision Making*, edited by Derek Koehler and Nigel Harvey, 89–109. Hoboken: Blackwell Publishing.

- Knoch, Utte. 2011. "Rating scales for diagnostic assessment of writing: What should they look like and where should the criteria come from?" *Assessing Writing* 16 (2): 81–96. <https://doi.org/10.1016/j.asw.2011.02.003>.
- Krosnick, Jon. 1991. "Response strategies for coping with the cognitive demands of attitude measures in surveys." *Applied Cognitive Psychology* 5 (3): 213–36. <https://doi.org/10.1002/acp.2350050305>.
- Kuiken, Folkert, and Ineke Vedder. 2017. "Functional adequacy in L2 writing: Towards a new rating scale." *Language Testing* 34(3): 321–36. <https://doi.org/10.1177/0265532216663991>.
- Laming, Donald. 2004. *Human Judgment. The Eye of the Beholder*. Stamford: Thomson Learning.
- Leder, Helmut, Benno Belke, Andriest Oeberst, and Dorothee Augustin. 2004. "A model of aesthetic appreciation and aesthetic judgements." *British Journal of Psychology* 95: 489–508. <https://doi.org/10.1348/0007126042369811>.
- Lengo, Nskala. 1995. "What is an error?" *English Teaching Forum* 33 (3): 20–24.
- Lennon, Paul. 1991. "Error: Some problems of definition, identification, and distinction." *Applied Linguistics* 12 (2): 180–96. <https://doi.org/10.1093/applin/12.2.180>.
- Lumley, Tom. 2005. *Assessing Second Language Writing: The Rater's Perspective*. Frankfurt: Lang.
- March, James. 1994. *A Primer on Decision-Making*. New York: Free Press.
- McHugh, Mary. 2012. "Interrater reliability: The Kappa statistic." *Biochemia Medica* 22 (3): 276–82. <https://doi.org/10.11613/BM.2012.031>.
- Milanovic, Michael, Nick Saville, and Shen Shuhong. 1996. "A study of the decision-making behaviour of composition markers." In *Performance Testing, Cognition and Assessment: Selected Papers from the 15th Language Testing Research Colloquium, Cambridge and Arnhem*, Michael Milanovic and Nick Saville, 92–114. Cambridge: UCLES and Cambridge University Press.
- Morgan, Candia. 1996. "The teacher as examiner: The case of mathematics coursework." *Assessment in Education: Principles, Policy and Practice* 3 (3): 353–75.
- Mumford, Simon, and Derin Atay. 2021. "Teachers' perspectives on the causes of rater discrepancy in an English for Academic Purposes context." *Assessing Writing* 48. <https://doi.org/10.1016/j.asw.2021.100527>.
- Mussweiler, Thoma, and Birte Englich. 2005. "Subliminal anchoring: Judgmental consequences and underlying mechanisms." *Organizational Behavior and Human Decision Processes* 98 (2): 133–43. <https://doi.org/10.1016/j.obhdp.2004.12.002>.
- Myers, Albert, Carolyn McConville, and William Coffman. 1966. "Simplex structure in the grading of essay tests." *Educational and Psychological Measurement* 26 (1): 41–54.
- Myford, Carol, and Edward Wolfe. 2003. "Detecting and measuring rater effects using many-facet Rasch measurement: Part I." *Journal of Applied Measurement* 4 (4): 386–422.
- Nagelkerke, Nico. 1991. "A note on a general definition of the coefficient of determination." *Biometrika* 78 (3): 691–92. <https://doi.org/10.1093/biomet/78.3.691>.
- Newell, Benjamin, David Lagnado and David Shanks. 2007. *Straight Choices: The Psychology of Decision Making*. New York: Psychology Press.
- Park, Yoon Soo, Jing Chen, and Seven Holtzman. 2015. "Evaluating efforts to minimize rater bias in scoring classroom observations." In *Designing Teacher Evaluation Systems*, edited by Thomas J. Kane, Kerri A. Kerr, and Robert C. Pianta, 381–414. Hoboken: John Wiley & Sons. <https://doi.org/10.1002/9781119210856.ch12>.
- Peters, Henry. 1942. "The experimental study of aesthetic judgments." *Psychological Bulletin* 39 (5): 273–305. <https://doi.org/10.1037/h0057008>.
- Postman, Leo, and Edward Tolman. 1959. "Brunswik's Probabilistic Functionalism." In *Psychology: A Study of a Science, Study 1: Conceptual and Systematic, Vol. 1: Sensory, Perceptual, and Physiological Formulations*, edited by Sigmund Koch, 502–64. New York: McGraw-Hill.
- Prall, David. 1929. *Aesthetic Judgment*. Georgia: Crowell.
- Quinlan, Thomas, Derrick Higgins and Susanne Wolff. 2009. "Evaluating the construct coverage of the e-rater® scoring engine." *ETS Research Report*.
- Richards, Jack C. 1974. *Error Analysis. Perspectives on Second Language Acquisition*. London: Longman.
- Rubinstein, Ariel. 1997. *Modeling Bounded Rationality*. Cambridge: MIT Press.

- Sadler, Royce. 1989. "Formative assessment and the design of instructional systems." *Instructional Science* 18: 119–44.
- Sakiy, Alfred. 2000. "Validation of holistic scoring for ESL writing assessment: How raters evaluate compositions." In *Fairness and Validation in Language Assessment*, edited by Antony John Kunnan, 129–52. Cambridge: Cambridge University Press.
- Scott, Susanne, and Reginald Bruce. 1995. "Decision-making style: The development and assessment of a new measure." *Educational and Psychological Measurement* 55 (5): 818–31.
- Shirazi, Masoumeh. 2012. "When raters talk, rubrics fall silent." *Language Testing in Asia* 2: 123–39.
- Sigott, Günther, Melanie Fleischhacker, Stephanie Sihler, and Jennifer Steiner. 2019. "The effect of written feedback types on students' academic texts: A pilot study." *AAA: Arbeiten aus Anglistik und Amerikanistik* 44 (2): 195–216.
- Sigott, Günther, Hermann Cesnik, and Nikola Dobrić. 2016. "Refining the Scope–Substance error taxonomy: A closer look at Substance." In *Corpora in Applied Linguistics – Current Approaches*, edited by Nikola Dobrić, Eva-Maria Graf, and Alexander Onysko, 79–95. Newcastle: Cambridge Scholars Publishing.
- Simon, Herbert. 1957. *Models of Man, Social and Rational: Mathematical Essays on Rational Human Behavior in a Social Setting*. New York: John Wiley and Sons.
- Skehan, Peter. 2009. "Modelling second language performance: Integrating complexity, accuracy, fluency, and lexis." *Applied Linguistics* 30 (4): 510–32. <https://doi.org/10.1093/applin/amp047>.
- Suto, Irenka, and Jackie Greatorex. 2008. "What goes through and examiner's mind? Using verbal protocols to gain insights into the GCSE marking process." *British Educational Research Journal* 34 (2): 213–33. <https://doi.org/10.1080/01411920701492050>.
- . 2006. "An empirical exploration of human judgement in the marking of school examinations." Paper presented at the 32nd Annual Conference of the International Association for Educational Assessment, Singapore.
- Swan, Michael, and Bernard Smith. 2001. *Learner English. A Teacher's Guide to Interference and Other Problems*. Cambridge: Cambridge University Press.
- Thunholm, Peter. 2004. "Decision-making style: Habit, style, or both?" *Personality and Individual Differences* 36 (4): 931–44. [https://doi.org/10.1016/S0191-8869\(03\)00162-4](https://doi.org/10.1016/S0191-8869(03)00162-4).
- Tigelaar, Dineke, Diana Dolmans, Ineke Wolfhagen, and Cees van der Vleuten. 2005. "Quality issues in judging portfolios: Implications for organizing teaching portfolio assessment procedures." *Studies in Higher Education* 30 (5): 595–610. <https://doi.org/10.1080/03075070500249302>.
- Tversky, Amos, and Daniel Kahneman 1992. "Advances in prospect theory: Cumulative representation of uncertainty." *Journal of Risk and Uncertainty* 5 (4): 297–323.
- van Someren, Maarten, Yvonne Barnard, and Jacobijn Sandberg. 1994. *The Think Aloud Method. A Practical Guide to Modelling Cognitive Processes*. London: Academic Press.
- Veall, Michael, and Klaus Zimmermann. 1996. "Pseudo-R2 measures for some common limited dependent variable models." *Sonderforschungsbereich 386: paper 18*. <https://doi.org/10.5282/ubm/epub.1421>.
- Wiliam, Dylan. 1992. "Some technical issues in assessment: A user's guide." *British Journal for Curriculum and Assessment* 2 (3): 11–20.
- Wolf, Alison. 1995. *Competence-Based Assessment*. Buckingham: Open University Press.
- Wolfe, Edward, and Tian Song. 2016. "Methods for monitoring and document rating quality." In *The Next Generation of Testing: Common Core Standards, Smarter-Balanced, PARCC, and the Nationwide Testing Movement*, edited by Hong Jiao and Robert Lissitz, 107–42. Information Age.
- Xie, Qin. 2019. "Error analysis and diagnosis of ESL linguistic accuracy: Construct specification and empirical validation." *Assessing Writing* 41: 47–62. <https://doi.org/10.1016/j.asw.2019.05.002>.
- Zhang, Jie. 2016. "Same text different processing? Exploring how raters' cognitive and meta-cognitive strategies influence rating accuracy in essay scoring." *Assessing Writing* 27: 37–53. <https://doi.org/10.1016/j.asw.2015.11.001>.

**Tania Tagle**, Temuco Catholic University  
**Paulo Etchegaray**, Autonomous University of Chile  
**Claudio Díaz**, University of Concepción  
**Mabel Ortiz**, Catholic University of the Most Holy  
 Concepción

**Marcela Quintana**, Arturo Prat University  
**Lucía Ramos**, Catholic University of the North, Chile

2024, Vol. 21 (1), 111-127(316)  
 journals.uni-lj.si/elope  
[https://doi.org/10.4312/](https://doi.org/10.4312/elope.21.1.111-127)  
 elope.21.1.111-127  
 UDC: 811.111'243:37.091.26(83)

## How Do In-Service EFL Teachers Assess Student Language Learning? Analysis of English Assessment Instruments Used in Chilean Secondary Schools

### ABSTRACT

The study analyses the assessment instruments designed by Chilean English as a Foreign Language (EFL) in-service teachers. The participants were 110 secondary school teachers from the central, northern, and southern areas of Chile. The data collection technique was document analysis of the assessment instruments designed in their teaching practices. The results suggest that the most typical instrument is a language test, while speaking/writing performance evaluations with rubrics and rating scales are used to a lesser degree, and all having writing, reading, grammar, and vocabulary as their primary assessment foci. It is suggested that universities and administrators of educational institutions promote opportunities for professional updating regarding language assessment. Moreover, pre-service teacher education needs to emphasize and strengthen this didactic dimension.

**Keywords:** language assessment, TEFL, in-service teachers, teacher professional development, pre-service teacher education

### Kako učitelji ATJ ocenjujejo učenje jezika? Analiza instrumentov za ocenjevanje znanja angleščine, ki se uporabljajo v čilskih srednjih šolah

### IZVLEČEK

Študija analizira ocenjevalne instrumente, ki so jih oblikovali čilski učitelji angleščine kot tujega jezika (ATJ). V raziskavi je sodelovalo 110 srednješolskih učiteljev ATJ iz osrednjega, severnega in južnega Čila. Tehnika zbiranja podatkov je bila analiza dokumentov z ocenjevalnimi instrumenti iz pedagoške prakse. Rezultati kažejo, da je najpogostejši instrument jezikovni test, medtem ko se ocenjevanje govorne ali pisne zmožnosti z rubrikami in ocenjevalnimi lestvicami uporablja v manjši meri, pri vseh pa so osnovna ocenjevalna področja pisanje, branje, slovnica in besedišče. Predlagamo, da univerze in vodje izobraževalnih ustanov spodbujajo priložnosti za strokovno izpopolnjevanje na področju ocenjevanja jezika. Poleg tega je treba pri začetnem izobraževanju učiteljev poudariti in okrepiti to didaktično razsežnost.

**Ključne besede:** ocenjevanje jezika, poučevanje angleščine kot tujega jezika, učitelji na delovnem mestu, strokovno izpopolnjevanje učiteljev, začetno izobraževanje učiteljev

# 1 Introduction

Chile is a country located in South America, whose official language is Spanish. At a national level, learning English as a Foreign Language (EFL) is prioritized to boost the country's competitiveness and accelerate its cultural and commercial integration with the rest of the globalized world (Government of Chile 2014). Moreover, the Chilean educational system considers EFL a mandatory school subject at primary education levels (from the fifth to the eighth grades, specifically) and during the four years of secondary education (Ministry of Education of Chile 2015). Based on the suggestions of the Chilean curriculum, students are expected to achieve a B1 level of English (Ministry of Education of Chile 2019). This means a lower-intermediate level according to the Common European Framework of Reference for Languages (CEFR) (Council of Europe 2001).

Nowadays, the practices of English language teaching and learning should promote the development of receptive skills (listening, reading) and productive skills (speaking, writing), following a communicative approach (Celce-Murcia 2014; Duff 2014; Kumaravadivelu 2006; Larsen-Freeman and Anderson 2011; Nunan 2015; Richards and Rogers 2014). The findings of studies specify that classroom language assessment practices can facilitate student achievements based on these linguistic competences (Namaziandost et al. 2020; Safdari and Fathi 2020). This is because English teachers collect the evidence of student learning through appropriate assessment instruments, which allows them to identify the impact of their instructional practices and, consequently, make the right decisions for evidence-based education.

Despite the above, studies show that many EFL teachers prioritize summative over formative assessment practices, focusing mainly on assessing linguistic knowledge rather than language skills (Minda and Chaka 2023; Tagle et al. 2022; Tsagari 2016; Wubshet and Menuta 2015; Yan, Zhang, and Dixon 2022). This behaviour might suggest that undergraduate professional training does not provide prospective teachers with sufficient language assessment literacy (Giraldo and Murcia 2019; Kong, Molnár, and Xu 2022). A study conducted by Tsagari and Vogt (2017) confirms this perspective, suggesting that EFL teachers' lack of assessment training results in their dependence on using published assessment materials or assuming the practices of others without questioning them. The researchers conclude that an insufficient theoretical basis for assessment practices can impede the implementation of assessment innovations.

In Chile, English teacher education programmes usually include a one-semester course on evaluating students from a general pedagogical perspective (British Council 2015). This university subject does not address the specific techniques and instruments for assessing foreign language skills. Moreover, it does not develop a reflective attitude towards the assessment process. Based on these facts, it has been demonstrated that many prospective teachers learn how to assess student performance during their professional internships rather than their pre-service teacher training (Earl 2012).

In 2017, Chilean third-year secondary education students took a mandatory national English language examination that evaluated their reading and listening skills. The results indicated that 68% of Chilean EFL learners had a beginner level of the foreign language (A1,

according to the CEFR), while 32% demonstrated basic to intermediate proficiency (A2 and B2) (Agencia de Calidad de la Educación 2017). The data showed that Chilean students struggle to develop their English language skills further. These results may thus shed light on how assessment processes are carried out and which aspects of foreign language teaching and learning are prioritized in Chile. This indicates that English language teaching and assessment practices may not adequately address the development of students' communicative skills at the national level.

Based on the background mentioned above, this research aims to analyse the assessment instruments designed by Chilean in-service EFL teachers in the context of their instructional practices. Likewise, the present study seeks to answer the following specific research questions:

- 1) Which kind of language assessment instruments do Chilean in-service EFL teachers design?
- 2) What are the language dimensions included in the language assessment instruments designed by Chilean in-service EFL teachers?
- 3) What are the characteristics of the assessment instruments designed by Chilean in-service EFL teachers?

## 2 Assessment in English Language Teaching and Learning

Educational assessment involves the systematic documentation of evidence of student performance against learning objectives (Britton 2021; Solano-Flores 2016). This process includes data collection and interpretation by teachers. Within this, assessment is not an isolated didactic dimension because it operates within a system that links its procedures with different elements, such as the curriculum, academic content, and standards of linguistic competence, in addition to the pedagogical process (Gottlieb 2016).

Assessment can be diagnostic, formative, and summative (Marsh 2010; Ravela, Picaroni, and Loureiro 2019). The first type is conducted at the beginning of a didactic unit to identify students' strengths and weaknesses to make decisions about lesson planning. Formative assessment is carried out during the pedagogical process to help students achieve their learning objectives. In contrast, summative assessment measures students' learning at the end of a didactic unit, generally using a grade.

Teachers' assessment practices can be associated with three approaches. First, it is possible to refer to the *assessment of learning*, which is summative in nature and is used to identify what students have learned as a result of a pedagogical intervention (Earl and Katz 2006). This type of assessment is usually implemented exclusively by the teacher to document learning, measuring and categorizing it to report this information to students, parents, and teachers (Chappuis and Stiggins 2020). In English language teaching and learning, these practices are linked to traditional assessment, encompassing instruments such as tests with multiple-choice, true-false, and gap-filling tasks (Harris 2013). Traditional assessment tends to focus on language accuracy, which means the correct use of linguistic structures or conceptual contents (Goh and Burns 2012). Research findings have posited that these evaluative practices seek, primarily, to prepare students to be successful test takers rather than to help them develop

communicative skills (Styron and Styron 2012; Wubshet and Menuta 2015). Moreover, other studies have reported that traditional assessment does not significantly impact the students' learning of English language skills (Forutan 2014; Goçtü 2012).

Another approach is *assessment for learning*, which is conducted for formative purposes. These assessment practices are employed to help students construct their learning throughout the instructional process (Butt 2010; Jones 2010). This involves providing continuous feedback and encouraging learners to self- and co-assess their performance. This type of assessment emphasizes communication or the production of clear ideas (fluency) over the correct use of linguistic structures (accuracy) (Türk 2022). Likewise, it understands learning as a social process so that interaction tasks and co-assessment/self-assessment procedures play a relevant role (Heritage 2022).

Alternatively, it is possible to refer to *assessment as learning*. This assessment approach corresponds to an active process of cognitive restructuring that occurs when learners interact with new ideas (Earl and Katz 2006). This type of assessment performance is related to metacognition, which implies knowledge of one's thinking processes.

## 2.1 English Assessment Instruments

Assessment instruments are used to collect information or evidence about students' learning (Chappuis and Stiggins 2020). Within the English language teaching and learning framework, instruments can be associated with traditional assessment, favouring the identification of linguistic knowledge, such as grammar, vocabulary, or pronunciation, and employing memorization (Forutan 2014; Shams and Tavakoli 2014). One of the most common instruments of this type of assessment is the written test, mainly employed in summative assessment practices as its purpose is to measure student language knowledge (Richard-Amato 2010).

It is also possible to refer to authentic assessment instruments, which focus on recognizing student learning concerning English language communicative competencies through formative assessment (Aliasin and Amianlu 2017; Brown and Abeywickrama 2018). These instruments relate to communicative performance assessment tasks performed in class, in which learners have an active role in listening, reading, speaking, and writing in English. These tasks are designed to help students achieve the learning objectives during the process, serving as evidence of achievement and allowing for feedback.

In authentic assessment or assessment for learning, checklists are used with instruments or tasks to gather information and judge language learners' performance (Gottlieb 2021). Checklists follow a yes/no format when evaluating whether students have met specific criteria by observing them (Katz 2014).

Teachers may use scales to supplement their assessment instruments when documenting students' language performance, especially in the case of productive skills (Fulcher and Davison 2007; Knoch 2009; Kuiken and Vedder 2017; Murray and Christison 2020). On the one hand, rating scales are used to establish the degree or frequency of students' displayed language functions, behaviours, skills, or strategies. These include mastery levels or bands

based on criteria. On the other hand, rubrics are criterion-referenced instruments utilized to evaluate learners' work based on uniform descriptors, which provide complete details on a specific performance. Rubrics can be holistic and analytic (Tedick and Lyster 2019). The former provides an overall assessment regarding student work, while the latter divides student performance into multiple dimensions or criteria.

Some authors suggest that rating scales and rubrics have distinct differences (see, for example, Plakans and Gebril 2015; Murray and Christison 2020). A rating scale is designed to determine a score, being more concise in describing different performance levels or indicators. In contrast, a rubric is mainly used to provide feedback to students, and therefore their criteria and descriptors must be more detailed and concrete.

## 2.2 Tasks Associated with the Assessment of Communicative Skills and Linguistic Content in English

In the context of English language learning, language assessment tasks collect information on students' achievement of learning objectives (Katz 2014). Regarding listening and reading, open-ended comprehension activities allow students to produce their own responses considering the general or specific ideas of an oral or written text (Oakhill, Cain, and Elbro 2015). It has also been proposed that the development of these skills can be assessed utilizing close-ended comprehension tasks, which allow students to respond only by selecting previously defined options. These tasks contain, for example, multiple-choice, true-false, or gap-filling items, short-answer questions, and matching, among others (Alderson 2000; Buck 2001).

In speaking assessment, open-ended and structured tasks are used. Luoma (2004) indicates that the former allow students to put the skill into practice in terms of language communicative functions, such as describing, narrating, giving instructions, and explaining, among others. The latter are not based on the learners' creation, but on the mastery of certain language components. Some examples of these types of activities include sentence drilling, reading aloud, oral sentence gap-filling exercises, and the formulation of short answer questions, among others.

To assess the learning objectives associated with writing in English, Brown and Abeywickrama (2018) highlight imitative tasks, which focus on the spelling correction of letters, words, or sentences. Likewise, the authors identify intensive or controlled tasks that involve the correct application of grammar or vocabulary when writing. Open-ended tasks, based on summarizing and integrating sentences within paragraphs, are also suggested. Finally, the authors refer to extensive written production tasks, in which learners must be competent in the ability to produce longer texts, such as essays, reports, and theses, among others.

## 3 Research Methodology

This study employs quantitative and qualitative research. The research design includes a collective case study, whose general objective is to analyse the assessment instruments designed by Chilean in-service EFL teachers. This research aim encompasses the types of assessment instruments, their respective assessment foci, and the tasks associated with these instruments.

### 3.1 Participants

The participants in this research were 110 EFL teachers working in secondary schools located in Chile. These institutions are located in the central, northern, and southern areas of the country. All study subjects signed a letter of informed consent, which specified that their participation was voluntary and that, in addition, the information they provided would be analysed anonymously, and their identity would be protected during and after the process.

### 3.2 Data Collection Technique

The document analysis technique was employed. According to Rapley (2018), this procedure involves examining materials produced by research participants to recognize their vision, actions, or what happens in their settings. In the present study, document analysis was used to process 824 assessment materials designed by the informants regarding their face-to-face instructional practices. Each of the teachers provided, voluntarily, between one and 10 documents.

First, quantitative document analysis was used by reviewing each of the instruments and completing a data matrix framed by the types of instruments, their assessment focus/foci (language dimension), and the tasks they are associated with. Then, the frequency of these components was calculated.

The ATLAS.ti software was used to interpret the instruments during the qualitative document analysis. The textual data were coded and subsequently organized into categories and subcategories, which focused on the characteristics of the documents. This was done by establishing relationships of meaning between the previously recognized codes. Conceptual networks were then created, centred on the findings.

## 4 Research Results

### 4.1 Assessment Instruments Designed by Chilean In-Service EFL Teachers

The quantitative document analysis revealed the frequency of the use of the 824 assessment instruments developed by the Chilean in-service EFL teachers in their instructional practices (see Table 1). Within this framework, the instrument they use the most is the *test* (74%). Then, in a smaller proportion of cases, these professionals design *speaking/writing performance evaluations with rubrics* (16%) and *speaking/writing performance evaluations with rating scales* (10%).

TABLE 1. Frequency of assessment instruments designed by in-service EFL teachers.

Assessment instrument	Frequency	Percentage
Test	610	74%
Speaking/writing performance evaluation with rubric	130	16%
Speaking/writing performance evaluation with rating scale	84	10%
<b>Total</b>	<b>824</b>	<b>100%</b>

Concerning the *speaking/writing performance evaluations*, these include scales, such as rubrics or rating scales. As stated in the Theoretical Framework, both help teachers evaluate, measure, and identify learners' language productive performance as they integrate criteria (aspects to assess/evaluate) and indicators (criteria's performance level). In Chile, rubrics provide students with more detailed and specific feedback, so they include descriptors (criteria's performance level) that are useful to determine the quality of learner production. On the other hand, rating scales assign a score to students, presenting indicators (performance levels) based on criteria without describing them.

## 4.2 Language Dimensions Encompassed in the Assessment Instruments Designed by In-Service EFL Teachers

Considering the document analysis of the assessment instruments designed by the Chilean in-service EFL teachers, it is relevant to establish the frequency of particular language dimensions. This involved calculating how many language skills (listening, reading, speaking, and writing) and/or linguistic systems (grammar, vocabulary, or pronunciation) each assessment instrument covered as a macro category.

Assuming that assessment instruments may evaluate one or multiple skills or linguistic systems in second/foreign language assessment, a value of 1 was given for the presence of each. After analysing the 824 documents, all the values for each dimension/system were summed to calculate their frequency and percentage.

Firstly, the instruments focus on the assessment of language skills (see Table 2). Within this framework, the language ability they most seek to assess is *writing* in English (39%), followed by *reading* (31.3%). Consequently, the instruments assess the development of *listening* (18.2%) and *speaking* (11.5%) skills to a lesser degree.

TABLE 2. Frequency regarding the consideration of language skills in the assessment instruments.

Language dimension	Dimension frequency	Percentage
Writing	574	39%
Reading	462	31.3%
Listening	269	18.2%
Speaking	169	11.5%
<b>Total</b>	<b>1,474</b>	<b>100%</b>

TABLE 3. Frequency regarding the consideration of linguistic contents in the assessment instruments.

Language dimension	Dimension frequency	Percentage
Vocabulary	726	50%
Grammar	654	45%
Pronunciation	76	5%
<b>Total</b>	<b>1,456</b>	<b>100%</b>

Secondly, in terms of the linguistic contents in English that the assessment instruments present (see Table 3), it is possible to observe that these tools are mostly based on *vocabulary* (50%) and *grammar* (45%). Finally, the instruments also show that *pronunciation* in English is assessed to a lesser degree (5%).

### 4.3 Characteristics of the Assessment Instruments Designed by In-Service EFL Teachers and the Frequency of Their Assessment Tasks

#### Test

The qualitative document analysis identified the *tests* as assessment instruments designed by the in-service EFL teachers, whose characteristics are illustrated through a conceptual network (see Figure 1). The types of test tasks included in these instruments relate to the following learning objectives: *to develop comprehension skills*, *to develop comprehension production skills*, and *to apply linguistic contents*.

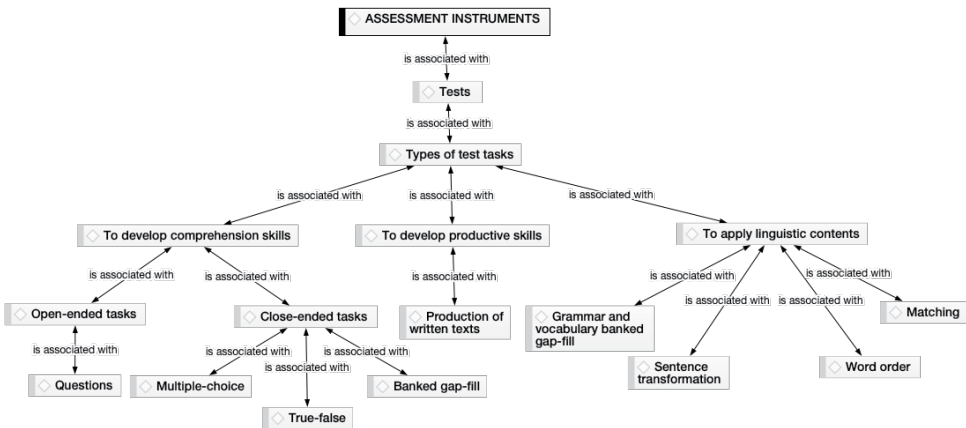


FIGURE 1. Tests as assessment instruments designed by in-service EFL teachers.

On the one hand, the *tests* designed by in-service EFL teachers include test tasks that seek to assess the learning objective *to develop comprehension skills*. The tests include reading activities (739 out of 1585, 46.6%) and listening activities (314 out of 1585, 19.8%), which are used similarly. They consist of *open-ended* and *close-ended tasks*, through which students must recognize and interpret general and specific information from oral or written texts that provide exposure to the foreign language. Regarding the document analysis, the *open-ended tasks* (187, 11.8% of the total) are mainly questions learners must answer with short or paragraph responses in English based on a text. The following test excerpt is an example of this type of reading test task: “Answer the following questions: What is the most serious problem presented in the text? What is the most effective plan of action? Why?” (Participant 14). Another participant formulates a similar listening test task: “What is the purpose of the text? Why?” (Participant 32).

On the other hand, within the framework of the learning objective *to develop comprehension skills*, the analysed tests contain items only related to *close-ended tasks* (866, 54.6% of the

total). In these the students must respond by producing or selecting a predetermined answer considering general and specific information from the texts they have read or listened to. Within this, the English tests show *multiple-choice test tasks* (283, 17.9% of the total), in which learners must answer, opting for only one of the choices presented to them. Of all the activities, 11.8% involve reading (187) and 6.1% involve listening (96). The following test excerpt is an example of this: “Listen again and answer by marking the correct alternative: What activity will Ted do? a) He will have a party, b) He will meet with his friends, c) He will study for his exams” (Participant 2). Another informant designed a similar item in one of his reading tests: “Regarding the written text, answer the questions, circling the correct alternative” (Participant 7).

Considering the listening or reading *close-ended tasks* included in English tests, another activity corresponds to *true-false tasks* (110, 6.9% of the total). In these, students must read statements presented to them and specify whether they integrate information that appears in the texts or not. Of all the cases, 5.8% involve reading (92), while only 1.1% involve listening (18). The following test excerpt refers to this type of reading test task: “According to the text you read, write T if the information is true or F if it is false: 1) The scene takes place in the living room; 2) Bob is visiting his family” (Participant 10). Another participant designed a similar item to assess listening: “While you are listening to the text, decide whether the following sentences are true or false. Justify the false ones” (Participant 35).

Based on this document analysis, *banked gap-fill tasks* are also *close-ended* exercises within tests that assess comprehension skills, worth 18.9% of the total test tasks (299). In these, students had to fill in blanks in paragraphs or sentences with specific information from oral and/or written input texts in English. Of all these exercises, 12.7% involve reading (201), while 6.2% involve listening (98). This can be seen in the following segment of a reading test task: “Complete these sentences with information from the text you are reading: a) Teenagers are different from those of the past. Their \_\_\_ habits, way of life, and \_\_\_ are an example” (Participant 12). Another test includes the following task: “Listen and fill in the blanks: 1) Cynthia’s hair is \_\_\_ and \_\_\_. She goes to the hairdresser every \_\_\_” (Participant 41).

Additionally, the English tests designed by the study subjects contain test tasks related to the learning objective *to develop production skills*, which is mostly associated with the writing skills in the foreign language (201 test tasks, worth 12.7% of the total). This involves the *production of written texts*, including from paragraphs to texts of greater complexity and length within a test. This is shown in the following test segment: “Write, in a paragraph, about your opinion based on the text you read. Justify your answer” (Participant 72). Another test task presents a similar item: “According to the text you listened to, write a short essay about pollution in the world. This should have an introduction, body, and conclusion” (Participant 98). It should be noted that this type of test task is graded by using rating scales or rubrics, which are part of the main evaluation instruments designed by the in-service EFL teachers.

Other English test tasks relate to the learning objective *to apply linguistic contents*. Through these, students must demonstrate their conceptual knowledge regarding the correct use of grammar and vocabulary. One of these test tasks focuses on *banked gap-filling* (136 test tasks, worth 8.6% of the total), in which learners must fill in the gaps of the provided sentences or

paragraphs, considering the relevant language structures or words. Of the total tasks, 4.9% are related to grammar (78), while only 3.7% are related to vocabulary (58). The following is an example: “Complete with don’t or doesn’t. Bill \_\_\_ plays tennis every Saturday” (Participant 11). Another text excerpt presents the instructions for a task related to vocabulary: “Write the following words in the empty spaces of the sentences: apply – temporary job – CV – interesting – full-time job – boring” (Participant 14).

*Sentence transformation* (6.4%, 101) is another test task associated with the learning objective *to apply linguistic contents*. In this context, students must paraphrase, demonstrating their knowledge of English grammar. The following test excerpt illustrates this form of assessment: “Considering the learning contents that relate to reported speech, create the second sentence, expressing a meaning similar to the first one and use the specified word” (Participant 20). Another test presents a similar item: “Transform these sentences, using the content of passive voice correctly: 1) Paul and Peter watched the soccer game” (Participant 54).

*Word order* (1.9%, 30) is also one of the types of test tasks present in English tests that evaluate the learning objective *to apply linguistic contents*. In these, students are presented with sentences whose components appear out of order. Therefore, they must rewrite them, considering their correct structure in the English language, as shown in the following example: “Order the following sentences: 1) I / the piano / played” (Participant 6). The following excerpt also integrates an equivalent test task: “Order the statements... 5) exciting / is / hobby / kayaking / an” (Participant 8).

There are also *matching* tasks based on lexical items in the tests (4.1%, 65), framed within the learning objective *to apply linguistic contents*. They essentially emphasize English vocabulary knowledge and include activities where students must match two equivalent test tasks: “Match the following words with their corresponding English definition” (Participant 14). A similar test also integrates one of these test tasks: “By writing an arrow, match the words in column A with their synonyms in column B” (Participant 28).

Table 4 refers to the frequency of tasks included in tests designed by the in-service EFL teachers. In this context, the quantitative document analysis helps to reveal that the most common tasks focus on assessing the development of *reading* (739, 46.6% of the tasks), *listening* (314, 19.8% of the tasks), *grammar* (209, 13.2% of the tasks), *writing* (201, 12.7%), and *vocabulary* (123, 7.8% of the tasks). Hence, it is possible to establish that in-service EFL teachers measure receptive skills, writing, and the application of linguistic contents to a greater extent by using tests.

The test task frequency was calculated by counting the assignments that were part of the tests provided by the participants. These instruments included from one to 10 test tasks with items within them.

### ***Speaking and writing performance evaluations utilizing rubrics and rating scales***

To assess the learning objective *to develop production skills*, the in-service EFL teachers also consider *speaking and writing performance evaluations* that are utilized with *rubrics* and *rating*

TABLE 4. Frequency of test tasks.

Test task	Task frequency	Percentage
Reading banked gap-fill	201	12.7%
Reading multiple-choice	187	11.8%
Answering reading questions	149	9.4%
Reading matching	110	6.9%
Sentence writing	103	6.5%
Grammar transformation	101	6.4%
Listening banked gap-fill	98	6.2%
Text writing	97	6.1%
Listening multiple-choice	96	6.1%
Reading true-false	92	5.8%
Grammar banked gap-fill	78	4.9%
Vocabulary matching	65	4.1%
Listening matching	64	4.0%
Vocabulary banked gap-fill	58	3.7%
Answering listening questions	38	2.4%
Word order	30	1.9%
Listening true-false	18	1.1%
<b>Total</b>	<b>1,585</b>	<b>100%</b>

*scales*. However, these are used to a lesser degree. Their characteristics are illustrated through a conceptual network (see Figure 2), and these instruments include assessment criteria related to language *fluency* and *accuracy*.

The analysed *rating scales* and *rubrics* contain assessment criteria, whose fulfilment is assessed through indicators that provide information about each level of performance (criteria) achieved by students. These are equivalent to scores. As stated in the Theoretical Framework, rubrics have more detailed descriptors (criteria's performance level), while rating scales present only the level of performance (indicator) for each assessment criterion. The analysed rubrics are all analytic, meaning they divide student performance into multiple dimensions or criteria.

Both scales are designed to assess the development of production skills in English to support *speaking and writing performance evaluations*. They incorporate, in the first place, *criteria* linked to *fluency*. Within this, students must demonstrate that their ideas are understandable and follow a logical order when writing or speaking in the foreign language. This is expressed in the following segment of a rating scale's criterion: "All interactions, in the role-playing performance, follow a clear sequence" (Participant 16). A similar element is evident in a rubric descriptor associated with an assessment criterion: "The student produces the language without interruptions. His/her contributions are relevant, and there is a clear organization of ideas (level 5)" (Participant 20).

On the other hand, the analysed *rating scales* and *rubrics* also have *assessment criteria* framed in language *accuracy*. They state that students must demonstrate the correct use of the

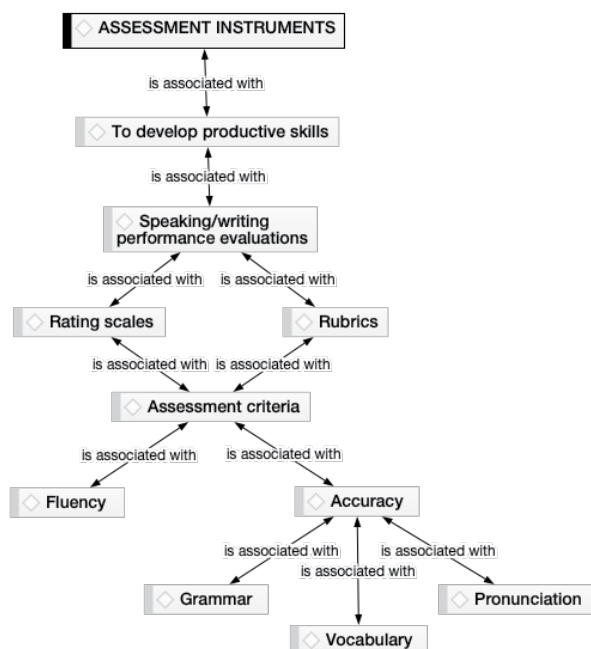


FIGURE 2. Speaking and writing performance evaluations utilized with rubrics and rating scales as assessment instruments designed by in-service EFL teachers.

conceptual knowledge of language form or lexicon when writing or speaking in English. For example, a segment of a rating scale refers to language accuracy: “The student, in his/her written report, writes correctly using the simple present tense in English” (Participant 32). Similarly, the following rubric descriptor exhibits the assessment of grammar knowledge: “Excellent performance – 3 points: The student demonstrates full command of simple and complex grammatical forms when speaking” (Participant 36).

The *assessment criteria* of the *rating scales* and *rubrics* used to assess the language *accuracy* of speaking and writing skills also allude to the *vocabulary* knowledge of English. This is illustrated below in a rating scale segment: “The student, in his/her oral presentation, uses words studied in the third learning unit” (Participant 29). This is further evidenced in the assessment criteria of another rubric: “5 points – No errors considering vocabulary. The student presents a high command in this dimension” (Participant 40).

Considering the assessment criteria linked to linguistic *accuracy*, these are also associated with knowledge of *pronunciation*. Within this, the assessment seeks to recognize whether the student correctly articulates the sounds of the language concerning vocabulary when developing the speaking skills. The following example presents a rating scale’s assessment criterion: “In the oral presentation, all words are pronounced correctly” (Participant 37). This aspect is also mentioned in a criterion descriptor present in a rubric used to assess speaking: “Not achieved (0 points) – The student presents many pronunciation problems that interfere with the comprehension of the dialogue” (Participant 9).

Table 5 presents the *frequency of assessment tasks* related to the *speaking and writing performance evaluations designed by in-service EFL teachers*. These activities are employed along with *rubrics* and *rating scales*. The quantitative document analysis made it possible to identify that these activities mainly focus on producing *written texts* (60.1%) and *oral presentations* (29%). The data show that the study subjects tend to assess writing skills more often with these instruments.

The task frequency was calculated by counting the speaking and writing performance instruments provided by the participants. These documents included the instructions or procedures and the rating scale or rubric.

TABLE 5. Frequency of assessment tasks related to speaking and writing performance evaluations utilized with rubrics and rating scales.

Assessment task	Frequency	Percentage
Production of written texts	172	60.1%
Oral presentations	83	29%
Open-ended speaking tasks	21	7.3%
Role-play	10	3.5%
<b>Total</b>	<b>286</b>	<b>100%</b>

## 5 Discussion of Results

The participants demonstrated a strong tendency to design their tests as an English language assessment instrument. This preference is related to traditional assessment, which may minimize the development of competences or performances that are put into practice in daily life (Al-Nouh, Taqi, and Abdul-Kareem 2014). The reason for this evaluative practice could be due to the versatility of this instrument, since in practice it is possible to adjust it to the characteristics of various educational scenarios (Giraldo 2018; Scully 2017). As a result, summative evaluations predominate, which could mean there are fewer opportunities to provide continuous student support or feedback (Butt 2010).

Additionally, in-service EFL teachers mostly consider writing and reading skills as their assessment foci, and listening and speaking skills are less emphasized. These assessment practices are also in line with what occurs in the area of linguistic contents, since those associated with pronunciation are almost entirely neglected. Consequently, learning the oral dimension of the foreign language and its assessment could be underrated, which then may have the effect that students do not develop the four linguistic skills. Therefore, experts in second language teaching and learning suggest that assessment in this area should consider the component of authenticity (see, for example, Brown and Abeywickrama 2018; Coombe 2012). This refers to designing assessment practices and instruments that incorporate language as naturally as possible, based on tasks involving communicative actions from daily life.

The document analysis presented in this study establishes that EFL teachers design production performance evaluations using rubrics and rating scales to a lesser degree, and show a tendency to evaluate the learning of writing. These practices may lack a link between

comprehension and production language skills, as in contexts where communication is developed authentically and functionally (Aliasin and Amianlu 2017).

Based on the previous points, the preference for summative/traditional assessment instruments that also prioritize certain language skills over others, segregating some or treating them separately, can be risky for the development of learners' English language skills, because the foreign language would not be taught within the framework of communicative approaches. Instead, different specialists suggest that language learning, and its corresponding assessment, should emphasize the use of both comprehension and production skills, simulating meaningful situations where the language is used as in everyday life (Murray and Christison 2020; Richard-Amato 2010). Several authors also state that effective language instruction requires opportunities for both language exposure and production (see, for example, Gottlieb 2016; Nunan 2015).

The findings of this research highlight some needs related to the didactic knowledge of in-service teachers of English. For this reason, it is suggested that higher education institutions and school administrators generate instances of professional development improvement and disciplinary updating. Moreover, teachers should reflect on the impact of their assessment practices on their students' learning of English (Giraldo 2019). In this sense, action research could be used so that teachers assume the role of researchers to contribute practical improvements to the community through a self-reflective, critical, and inquiry-oriented process (see, for example, Burns 2010). The principles of this proposal perceive teachers as reflective professionals who evaluate their practices and professional growth, constantly questioning themselves about student learning.

In addition, English teacher education institutions need to strengthen and monitor the development of the assessment competency of pre-service teachers in the undergraduate subjects of didactics and disciplinary practice (Tagle et al. 2022). Therefore, potential English teachers should be familiar with a variety of language assessment instruments, both traditional and authentic, that focus on comprehension and productive language skills from a communicative perspective. Likewise, in their professional preparation pre-service EFL teachers should be encouraged to design and implement assessment practices by questioning and reflecting on their impact on student English language learning.

## 6 Conclusions

The findings of the present study indicate that the in-service teachers of English mainly design their tests as an evaluation instrument, considering production performance evaluations using rubrics and rating scales to a lesser extent. Their assessment practices focus on the writing and reading skills, and the linguistic contents of vocabulary and grammar, which suggests that the oral dimension of the language is considered a secondary aspect. This logic may hinder the learning of the four linguistic skills in an integrated manner, since they would be taught in a piecemeal fashion, not following the principles of communicative approaches.

As per the above, the results show a strong tendency to use traditional and summative evaluation, marginalizing authentic and formative assessment. Consequently, these

practices may contribute little to constructing and verifying student learning related to their communicative competences in a foreign language.

It is suggested that universities and administrators of educational institutions encourage in-service teachers of English to increase and update their didactic knowledge focused on assessment competencies. Similarly, teacher education programmes should strengthen courses focused on this didactic dimension, favouring the design and implementation of assessment instruments in practice, complementing this with the competency of pedagogical reflection. Within this, longitudinal studies should be conducted to monitor how the assessment practices of teaching professionals change from the time they are undergraduate students until they are integrated into educational establishments as in-service teachers. This would allow us to keep track of the impact of pre-service teacher training on the pedagogical performance of candidates.

## References

- Agencia de Calidad de la Educación. 2017. "Informe de Resultados Estudio Nacional Inglés III Medio." Accessed April 04, 2022. [http://archivos.agenciaeducacion.cl/Informe\\_Estudio\\_Nacional\\_Ingles\\_III.pdf](http://archivos.agenciaeducacion.cl/Informe_Estudio_Nacional_Ingles_III.pdf).
- Al-Nouh, Nowreya A., Hanan A. Taqi, and Muneera M. Abdul-Kareem. 2014. "EFL primary school teachers' attitudes, knowledge and skills in alternative assessment." *International Education Studies* 7 (5): 68–84. <https://doi.org/10.5539/ies.v7n5p68>.
- Alderson, J. Charles. 2000. *Assessing Reading*. Cambridge: Cambridge University Press.
- Aliasian, Seyed Hesamuudin, and Mehdi Amianlu. 2017. "The effect of alternative assessment techniques on EFL learners' reading comprehension ability and self-efficacy in reading: The case of Iranian junior high school students." *Linguistics and Literature Studies* 5 (3): 160–68. <https://doi.org/10.13189/lis.2017.050302>.
- British council. 2015. "Annual report 2014–15." Accessed April 05, 2022. <https://www.britishcouncil.org/sites/default/files/annual-report-2014-2015.pdf>.
- Britton, Maria. 2021. *Assessment for Learning in Primary Language Learning and Teaching*. Bristol: Multilingual Matters.
- Brown, H. Douglas, and Priyanvada Abeywickrama. 2018. *Language Assessment: Principles and Classroom Practices*. 3rd ed. Thousand Oaks: Pearson.
- Buck, Gary. 2001. *Assessing Listening*. Cambridge: Cambridge University Press.
- Burns, Anne. 2010. *Doing Action Research in English Language Teaching: A Guide for Practitioners*. New York: Routledge.
- Butt, Graham. 2010. *Making Assessment Matter*. London: Continuum.
- Celce-Murcia, Marianne. 2014. "An overview of language teaching methods and approaches." In *Teaching English as a Second or Foreign Language*, edited by Marianne Celce-Murcia, Donna M. Brinton and Marguerite Ann Snow, 4th ed., 15–30. Boston: National Geographic Learning.
- Chappuis, Jan, and Rick Stiggins. 2020. *Classroom Assessment for Student Learning: Doing It Right – Using It Well*. 3rd ed. New York: Pearson.
- Coombe, Christine. 2012. "Second language assessment." In *The Cambridge Guide to Second Language Assessment*, edited by Christine Coombe, Peter Davidson, Barry O'Sullivan, and Stephen Stoyanoff, 1–6. New York: Cambridge University Press.
- Council of Europe. 2001. *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*. Cambridge: Cambridge University Press.
- Duff, Patricia A. 2014. "Communicative language teaching." In *Teaching English as a Second or Foreign Language*, 4th ed., edited by Marianne Celce-Murcia, Donna M. Brinton and Marguerite Ann Snow, 15–30. Boston: National Geographic Learning.

- Earl, Lorna M. 2012. *Assessment As Learning: Using Classroom Assessment to Maximize Student Learning*. 2nd ed. Thousand Oaks: Corwin.
- Earl, Lorna M., and Steven Katz. 2006. *Rethinking Classroom Assessment with Purpose in Mind*. Winnipeg: Manitoba Education, Citizenship and Youth Cataloguing in Publication Data.
- Forutan, Ali. 2014. "Traditional versus alternative writing assessment." *International Journal of Foreign Language Teaching & Research* 2 (7): 10–22.
- Fulcher, Glenn, and Fred Davidson. 2007. *Language Testing and Assessment: An Advanced Resource Book*. Abingdon: Routledge.
- Giraldo, Frank. 2018. "Diagnostic study on teachers' beliefs and practices in foreign language assessment." *Íkala, Revista de Lenguaje y Cultura* 23 (1): 25–44. <https://doi.org/10.17533/udea.ikala.v23n01a04>.
- . 2019. "Language assessment practices and beliefs: Implications for language assessment literacy." *HOW* 26 (1): 35–61. <https://doi.org/10.19183/how.26.1.481>.
- Giraldo, Frank, and Daniel Murcia. 2019. "Language assessment literacy and the professional development of pre-service language teachers." *Colombian Applied Linguistics Journal* 21 (2): 243–59. <https://doi.org/10.14483/22487085.14514>.
- Goçtü, Ramazan. 2012. "Comparison of traditional and portfolio assessment efficiency in English language teaching in high schools." *Journal of Education* 1 (2): 43–52.
- Goh, Christine C. M., and Anne Burns. 2012. *Teaching Speaking: A Holistic Approach*. New York: Cambridge University Press.
- Gottlieb, Margo. 2016. *Assessing English Language Learners: Bridges to Educational Equity*. 2nd ed. Thousand Oaks: Corwin.
- . 2021. *Classroom Assessment in Multiple Languages: A Handbook for Teachers*. Thousand Oaks: Corwin.
- Government of Chile. 2014. "National English Strategy 2014-2030." Accessed April 04, 2022. <https://www.economia.gob.cl/wp-content/uploads/2014/03/140307-Documento-Estrategia-Nacional-de-Ingl%C3%A9s-2014-2030.pdf>.
- Heritage, Margaret. 2022. *Formative Assessment: Making It Happen in the Classroom*. 2nd ed. Thousand Oaks: Corwin.
- Jones, Jane. 2010. "The role of assessment for learning in the management of primary to secondary transition: Implications for language teachers." *Language Learning Journal* 35 (2): 175–91. <https://doi.org/10.1080/09571730902928052>.
- Katz, Anne. 2014. "Assessment in second language classrooms." In *Teaching English as a Second or Foreign Language*, edited by Marianne Celce-Murcia, Donna M. Brinton, and Marguerite Ann Snow, 320–29. Boston: National Geographic Learning.
- Knoch, Ute. 2009. *Diagnostic Writing Assessment: The Development and Validation of a Rating Scale*. Frankfurt am Main: Peter Lang.
- Kong, Yunjun, Edit Katalin Molnár, and Nan Xu. 2022. "Pre- and in-service teachers' assessment and feedback in EFL writing: Changes and challenges." *SAGE Open* 12 (3). <https://doi.org/10.1177/2158244022112667>.
- Kuiken, Folkert, and Ineke Vedder. 2017. "Functional adequacy in L2 writing: Towards a new rating scale." *Language Testing* 34 (3): 321–36. <https://doi.org/10.1177/0265532216663991>.
- Kumaravadivelu, Bala. 2006. "TESOL methods: Changing tracks, challenging trends." *TESOL Quarterly* 40 (1): 59–81. <https://doi.org/10.2307/40264511>.
- Larsen-Freeman, Diane, and Marti Anderson. 2011. *Techniques and Principles in Language Teaching*. 3rd ed. Oxford: Oxford University Press.
- Luoma, Sari. 2004. *Assessing Speaking*. Cambridge: Cambridge University Press.
- Marsh, Colin. 2010. *Becoming a Teacher: Knowledge, Skills and Issues*. 5th ed. Frenchs Forest: Pearson.
- Minda, Motuma Hirpassa, and Chaka Chaka. 2023. "Implementing alternative assessment in a communicative English skills course at three Ethiopian universities." *MEXTESOL Journal* 47 (2): 1–14.
- Ministry of Education of Chile. 2015. *Bases Curriculares: 7º Básico a 2º Medio*. Santiago: Ministry of Education of Chile.
- . 2019. *Bases Curriculares: 3º a 4º Medio*. Santiago: Ministry of Education of Chile.

- Murray, Denise E., and MaryAnn Christison. 2020. *What English Language Teachers Need to Know: Facilitating Learning*. 2nd ed. New York: Routledge.
- Namaziandost, Ehsan, Samir Alekasir, Murad Hassan Mohammed Sawalmeh, and M. Zaini Miftah. 2020. "Investigating the Iranian EFL learners' attitudes towards the implementation of e-portfolios in English learning and assessment." *Cogent Education* 7 (1). <https://doi.org/10.1080/2331186X.2020.1856764>.
- Nunan, David. 2015. *Teaching English to Speakers of Other Languages: An Introduction*. New York: Routledge.
- Oakhill, Jane, Kate Cain, and Carsten Elbro. 2015. *Understanding and Teaching Reading Comprehension: A Handbook*. Abingdon: Routledge.
- Plakans, Lia, and Atta Gebril. 2015. *Assessment Myths: Applying Second Language Research to Classroom Teaching*. Ann Arbor: University of Michigan Press.
- Rapley, Tim. 2018. *Doing Conversation, Discourse and Document Analysis*. 2nd ed. London: SAGE.
- Ravela, Pedro, Beatriz Picaroni, and Graciela Loureiro. 2019. *¿Cómo Mejorar la Evaluación en el Aula? Reflexiones y Propuestas de Trabajo para Docentes*. 2nd ed. Montevideo: Grupo Magro Editores.
- Richard-Amato, Patricia A. 2010. *Making it Happen: From Interactive to Participatory Language Teaching*. 4th ed. White Plains: Pearson Education.
- Richards, Jack C., and Theodore S. Rodgers. 2014. *Approaches and Methods in Language Teaching*. 3rd ed. Cambridge: Cambridge University Press.
- Safdari, Maryam, and Jalil Fathi. 2020. "Investigating the role of dynamic assessment on speaking accuracy and fluency of pre-intermediate EFL learners." *Cogent Education* 7 (1): 1–19. <https://doi.org/10.1080/2331186X.2020.1818924>.
- Scully, Darina. 2017. "Constructing multiple-choice items to measure higher-order thinking." *Practical Assessment, Research and Evaluation* 22 (4): 1–13. <https://doi.org/10.7275/swgt-rj52>.
- Shams, Nasrin, and Mansoor Tavakoli. 2014. "The effect of peer, self, and traditional assessment on Iranian EFL learners' L2 reading comprehension." *Journal of Applied Linguistics and Language Research* 1 (1): 29–44.
- Solano-Flores, Guillermo. 2016. *Assessing English Language Learners: Theory and Practice*. New York: Routledge.
- Styron, Jennifer L., and Ronald A Styron. 2012. "Teaching to the test: A controversial issue in quantitative measurement." *Systemics, Cybernetics and Informatics* 10 (5): 22–25. <https://doi.org/10.54808/JSCI>.
- Tagle, Tania, Paulo Etchegaray, Claudio Díaz, Paola Alarcón, and Mabel Ortiz. 2022. "Representaciones Metafóricas sobre la Evaluación en el Aprendizaje y la Enseñanza del Inglés que poseen Estudiantes de Pedagogía y Profesores Noveles Chilenos." *Pensamiento Educativo, Revista de Investigación Latinoamericana (PEL)* 59 (2): 1–16. <https://doi.org/10.7764/PEL.59.2.2022.12>.
- Tedick, Diane J., and Roy Lyster. 2019. *Scaffolding Language Development in Immersion and Dual Language Classrooms*. Abingdon: Routledge.
- Tsagari, Dina. 2016. "Assessment orientations of state primary EFL teachers in two Mediterranean countries." *CEPS Journal* 6 (1): 9–30. <https://doi.org/10.26529/cepsj.102>.
- Tsagari, Dina, and Karin Vogt. 2017. "Assessment literacy of foreign language teachers around Europe: Research, challenges and future prospects." *Papers in Language Testing and Assessment* 6 (1): 41–63.
- Türk, Faruk. 2022. *Developing Oral Fluency in English*. Bursa: Sentez Publishing.
- Wubshet, Hailu, and Fekede Menueta. 2015. "Investigating the practice of alternative assessment in English classrooms: The case of selected grade nine English teachers' assessment practices." *International Journal of Scientific Research in Education* 8 (4): 159–71.
- Yan, Qiaozhen, Lawrence Jun Zhang, and Helen R. Dixon. 2022. "Exploring classroom-based assessment for young EFL learners in the Chinese context: Teachers' beliefs and practices." *Frontiers in Psychology* 13 (November). <https://doi.org/10.3389/fpsyg.2022.1051728>.



**Lena Tica**

University of Kragujevac, Serbia

**Ivana Krsmanović**

University of Kragujevac, Serbia

2024, Vol. 21 (1), 129-149(316)

journals.uni-lj.si/elope

<https://doi.org/10.4312/elope.21.1.129-149>

UDC: 811.111'276:004.8

# Overcoming the Writer's Block? Exploring Students' Motivation and Perspectives on Using ChatGPT as a Writing Assistance Tool in ESP

## ABSTRACT

Employing a mixed-method design (Pearson correlation, ANOVA analyses, and thematic analysis), the study explores the complex interplay between student motivations, English proficiency levels, and the perceptions of ChatGPT as a writing assistance tool. The results reveal a lack of significant correlation between overall motivation to learn English and ChatGPT usage, suggesting independence between motivational levels and technology adoption. The respondents (N=79) expressed polarized attitudes, with a majority praising ChatGPT's versatility, speed, and accuracy, while a substantial number voiced concerns about its potential limitations. Finally, the study identifies that the most developed instrumental motivation aligns with increased ChatGPT usage, underscoring its practical appeal by which ChatGPT is perceived as a handy tool that can complement the respondents' instrumental motivation to achieve their professional goals. The findings contribute to a deeper understanding of user experiences with AI-driven writing tools in language learning contexts.

**Keywords:** computer-mediated communication, English for specific purposes, artificial intelligence, ChatGPT, ELT

## Kako premagati krč pri pisanju? Motivacija in stališča študentov o uporabi ChatGPT kot pripomočka pri pisanju v angleščini za posebne namene

### IZVLEČEK

V prispevku je s pomočjo mešane metode (Pearsonova korelacija, analiza ANOVA in tematska analiza) prikazan zapleten medsebojni odnos med motivacijo učencev, stopnjo znanja angleščine in dojemanjem ChatGPT kot orodja za pomoč pri pisanju. Rezultati niso pokazali signifikantne povezave med splošno motivacijo za učenje angleščine in uporabo ChatGPT, kar kaže na neodvisnost med ravnmi motivacije in sprejemanjem tehnologije. Anketiranci (N=79) so imeli zelo različna stališča, saj jih je večina pohvalila vsestranskost, hitrost in natančnost ChatGPT, medtem ko so mnogi izrazili zaskrbljenost zaradi njegovih morebitnih omejitev. Raziskava tudi kaže, da se najvišja stopnja motivacije ujema s pogostejšo uporabo ChatGPT, kar poudarja njegovo praktično privlačnost, saj ChatGPT dojemajo kot priročno orodje, ki lahko dopolni motivacijo za doseganje poklicnih ciljev. Ugotovitve prispevajo k boljšemu razumevanju izkušenj uporabnikov z orodji za pisanje, ki jih poganja umetna inteligenca pri učenju jezikov.

**Gljučne besede:** računalniško posredovano sporazumevanje, angleščina za posebne namene, umetna inteligenca, ChatGPT, poučevanje angleščine

# 1 Introduction

The ever-evolving field of language learning is inherently dynamic, adapting to technological advances and societal shifts that, in turn, influence the methods through which we acquire and master languages. The role of technology in language learning has become so pronounced in the last two decades when students are born as “digital natives” (Vukićević-Đorđević 2015, 486), that the very notion of teaching and learning English without technology has become inconceivable. From the gamification of language learning and interactive language apps, to immersive virtual experiences, the pedagogical landscape has evolved to engage students in ways that were once considered far-fetched. The classroom is now no longer confined to four walls, but has become a dynamic space where technology serves as a catalyst for interactive and collaborative learning. The traditional chalk-and-talk approach has given way to a more student-centric and participatory model, fostering a deeper connection between learners and the English language.

It was against this backdrop of evolving educational methodologies and digital integration that late 2022 witnessed the introduction of ChatGPT (Generative Pre-trained Transformer), a revolutionary artificial intelligence (AI) language model. In a remarkably short span, ChatGPT permeated diverse domains, with English Language Teaching (ELT) standing out prominently in this dynamic transformation. As ChatGPT is predominantly designed for writing tasks, its introduction swiftly became a game-changer, foreshadowing a paradigmatic shift in how writing tasks are conceived, executed, and evaluated in ELT.

The initial response to this transformative tool unfolded on two fronts: it was marked by a palpable excitement, particularly from students, but simultaneously triggered apprehension among educators as they grappled with the potential challenges and ethical considerations that would accompany this development. Early headlines suggested an apocalyptic view of college essays (Marche 2022; Spatford 2023). However, this perspective gradually evolved, and newer articles now advocate for embracing, rather than banning, ChatGPT (Roose 2023; Byrd 2023; Claybourn 2023; Perez 2023). Rather than signifying a foreboding demise, ChatGPT holds immense promise to revolutionize ELT education, offering invaluable assistance and forging novel pathways for advancement in writing assignments (Smith 2023). Nevertheless, the conflicting opinions that still remain underscore the need for further research to thoroughly assess the potential and educational benefits of integrating ChatGPT in academic settings.

To comprehensively understand the impact, and employing a mixed-method design, this study explores university students’ perspectives on the utilization of ChatGPT as a writing assistance tool in English for Specific Purposes (ESP). Concurrently, the study investigates the relationship between students’ motivation to learn English and their engagement with this transformative tool, shedding light on the nuanced dynamics that underpin the integration of AI technology in language education.

## 2 Theoretical Background

Due to its complexity and unpredictable dynamics, motivation has long been a widely researched phenomenon, with the various perspectives on it changing dramatically over

time, undergoing scientific scrutiny and constantly bringing novel and fresh concepts to light. Although motivation was initially understood as a mere biological drive, and was later perceived to have more of a behavioural and/or cognitive nature, what remains indisputable (notwithstanding the numerous debates around it), is that motivation stands as one of the main determinants of second/foreign language (L2) learning achievement (Dörnyei 1994, 273; McDonough 1981, 142; Hardianti and Murtafi'ah 2022). Serving as an intrinsic or extrinsic driving force (Deci and Ryan 1985), motivation encourages students to engage with the language, persevere through challenges, and maintain an active and positive approach to the learning process.

Given that motivation stands out among the myriad factors influencing language acquisition, including age, intelligence, learning achievements, aptitudes, anxiety, personality, attitudes, self-identity, and parental encouragement, exploration of how the motivated actions of learners affect language learning outcomes has been an essential area of research in the field of linguistics and education. Gardner and Lambert (1972) were pioneers in emphasizing the central role of motivation in language learning, a perspective further strengthened by subsequent scholars such as Dörnyei (1990). Early recognition of the capricious nature of motivation came from McDonough (1981) who defined it as “[...] a dustbin – to include a number of possibly distinct concepts, each of which may have different origins and different effects and require different classroom treatment” (McDonough 1981, 143). Furthermore, Gardner’s socio-psychological approach to motivational theory, which perceives motivation as stemming from an interest in interacting and self-identification with the target language community (Crookes and Schmidt 1991, 470–71), is still considered the most influential approach to motivation in L2 learning and teaching. Based on Gardner’s socio-educational model, two basic types of motivation are “instrumental” and ‘integrative’ (Brown 2000, 162). Whereas an integratively motivated learner “wishes to identify himself with or become integrated into the society” that speaks the target language (Gardner 1983, 223) and assimilate with the target language community and culture, the instrumental motivation for acquiring a language is of a more utilitarian nature, seen as a desire to gain social recognition or economic advantages, best illustrated in a learner’s wish to learn a language so as to pass an exam, find a better-paid job, etc. Additionally, Cooper and Fishman (1977, 243) introduced a third dimension to this motivational typology, known as “personal” motivation. This dimension encompasses the intrinsic joy and personal satisfaction that learners derive from the process of learning a language. It signifies a motivation that goes beyond external rewards and speaks to the inherent gratification found in the pursuit of linguistic proficiency.

Among the four language skills – listening, speaking, reading, and writing – writing is widely acknowledged as one of the most challenging. Mastering the intricacies of grammar, sentence structure, and vocabulary, while simultaneously navigating cognitive, sociocultural, and linguistic competencies, places a considerable demand on learners. Numerous studies (Raimes 1983; Byrne 1988; Harmer 1998; Ellis 2015; Oliveira and Silva 2016; Ling 2016; Fareed, Ashraf, and Bilal 2016; Hyland 2003) have consistently recognized the critical importance of strong writing skills, particularly in the academic context. The ability to produce compelling, error-free pieces tailored for specific academic or professional purposes is a cornerstone of language proficiency. What is more, as Hedge argues (1988, 5), writing serves

as a supportive tool for learning, because it aids in reinforcing the understanding of newly acquired structures or vocabulary, and assists students in retaining information. The realm of researching the acquisition of writing skills naturally encompasses motivation, addressing the intricate challenges inherent in the writing process. Recent research, exemplified by Cahyono and Rahayu (2020) and Chen (2021), has delved into the correlation between students' motivation and their writing proficiency. These studies demonstrate that students with higher motivation levels exhibit increased engagement and success in writing tasks.

## 2.1 ChatGPT and ESP

Considering the inherent difficulty in mastering writing skills, it becomes reasonable to expect that AI language tools, including ChatGPT, would be fully embraced by ESP and ESL learners. ChatGPT, with its advanced language generation features, offers valuable assistance in overcoming challenges associated with writing (Tica, Krsmanović, and Lemper 2023). It serves as a supportive tool, providing learners with prompts, suggestions, and structuring assistance, thereby reinforcing the understanding of newly acquired language structures and vocabulary. Furthermore, the tool's versatility makes it suitable for a range of writing tasks, aligning with the diverse needs of learners.

Remarkably, within just a year of its introduction ChatGPT has attracted significant attention in academic research. A multitude of studies have explored its application across various domains, showcasing the tool's versatility. Notably, a substantial number of research endeavours have investigated its implications for English Language Teaching (ELT). This extensive exploration includes research into ChatGPT's potential and recommendations for use in teaching (Kostka and Toncelli 2023; Nugroho, Putro, and Syamsi 2023; Meniado 2023; Baskara 2023; Kohnke, Moorhouse, and Zou 2023), its impact on writing (Athanasopoulos et al. 2023; Fitria 2023; Zadorozhnyy and Lai 2024), as well as examinations of students' perspectives (Xiao and Zhi 2023) and teachers' views (Ulla, Perales, and Busbus 2023), and a significant number of other related studies that deal with the intersection of ChatGPT and ELT.

Despite the established link between motivation and writing proficiency, there exists a significant research gap in exploring the relationship between ChatGPT and motivation in English language learning, a domain that remains largely uncharted. Yildiz (2023) delved into the impact of integrating ChatGPT-generated dialogues into language teaching materials, specifically exploring its effect on the motivation of language learners. The results indicated that students in the experimental group, who employed ChatGPT, outperformed their counterparts in the post-tests and demonstrated higher motivation. Similarly, Ali et al. (2023) focused on the student perspective regarding motivation and the use of ChatGPT, employing a quantitative research approach. Their findings revealed that ChatGPT generally motivates learners to develop reading and writing skills, while the respondents expressed neutral attitudes towards its impact on developing listening and speaking skills. Taking these studies into consideration, this study attempts to address the indicated gap in the literature. By focusing exclusively on writing tasks and motivation, and employing both qualitative and quantitative methods, we aim to provide a more nuanced understanding of the impact of ChatGPT on language learners.

### 3 Methodology

#### 3.1 Research Context

The study aims to provide a comprehensive understanding of the multifaceted relationship between students' motivations, self-reported proficiency in English, attitudes towards ChatGPT, and its impact on their writing assignments in the context of an English language course for IT students. As for the treatment, at the commencement of the English language course designed for first-year IT students (during the summer term of the 2022/2023 academic year) the ESP instructors actively encouraged the use of ChatGPT as a writing assistance tool for different writing assignments, both within the classroom and at home. Providing explicit instructions on how to use ChatGPT prior to students' usage of the tool, the lecturers encouraged interactive demonstrations, fostering hands-on experience, with the goal of familiarizing students with the application. Importantly, it is worth noting that the use of ChatGPT was not forced, but instead the students were given the autonomy to decide when and how to incorporate it into their writing processes. As the semester progressed, the students kept handing in their written assignments while also deploying ChatGPT along the way. Additionally, whenever the students utilized ChatGPT, further discussions and comments on their assignments were provided to enhance the learning experience. Finally, the researchers' intention was to comprehensively explore students' views regarding the integration of ChatGPT into their writing processes, by administering a questionnaire at the end of the semester (the questionnaire is included in this paper's appendix). At the beginning of the questionnaire the participants were presented with a statement indicating that their participation was voluntary and that the data collected would be used solely for research purposes, with anonymity and confidentiality ensured.

More specifically, the study focuses on answering the following research questions:

- (1) What are the predominant types of motivation (instrumental, integrative, and personal) among university students, and how do students' motivations for learning English correlate with their frequency of using ChatGPT for writing assignments within the English language course for IT students during the 2022/2023 academic year?
- (2) To what extent does the proficiency level of students influence their attitudes towards ChatGPT as a writing assistance tool in an English language course for IT students during the 2022/2023 academic year?
- (3) What are the attitudes of university students towards the use of ChatGPT as a writing assistance tool, specifically in the context of completing homework assignments?

#### 3.2 Research Instrument and Sample

Employing a mixed-method design, our study engaged a sample of  $N=79$  university students enrolled in the English Language for IT course. The research instrument was developed by combining an adapted motivation scale (derived from Gardner, 1985, and Cooper and Fishman, 1977) and a set of original questions designed by the authors.

The instrument comprised three distinct parts: (I) demography of the sample (three questions); (II) motivation-related questions (seven questions); and (III) ChatGPT-related questions (seven questions). In the second section, a scale adapted from Al-Tamimi and Shuib (2009) was utilized, building upon Gardner's (1985) Attitude/Motivation Test Battery (AMTB) and Cooper and Fishman's (1977) scales, to assess three motivational constructs – instrumental motivation, integrative motivation, and personal motivation. An additional open-ended question was included for thematic analysis. The third section, comprising questions crafted by the researchers, explored the frequency of ChatGPT usage for English writing tasks, the types of tasks for which it was predominantly utilized, and an evaluation of ChatGPT's efficacy in English writing. Three yes/no questions were incorporated to assess the reliability and accuracy of ChatGPT's responses, participants' satisfaction, and perceived ease of access when using ChatGPT. An open-ended question sought to gather further insights into the participants' experiences by exploring the potential advantages and limitations of the tool. For the six questions in Part II, and three questions in Part III a five-point Likert scale (1 = strongly disagree, 2 = disagree, 3 = neutral, 4 = agree, 5 = strongly agree) provided the scoring framework, with means categorized as follows: 1.00–1.79 (very low); 1.80–2.59 (low); 2.60–3.39 (moderate); 3.40–4.19 (high); 4.20–5.00 (very high).

The reliability of the scale was assessed using the Cronbach's alpha coefficient, resulting in a Cronbach's alpha of 0.95 for the entire instrument, 0.732 for the motivation scale, and 0.96 for the ChatGPT-related questions in Part III.

Data were collected through the Google Forms tool. The analysis was conducted utilizing SPSS and ANOVA analysis tools, with Pearson coefficient analysis applied to determine the correlation between motivation and attitudes towards ChatGPT. Thematic analysis was employed to extract meaningful insights from the responses to open-ended questions.

In Section I of the questionnaire, the participants were asked to provide demographic information. The sample consisted of 72.5% males, 25% females, and 2.5% who chose not to disclose their gender. In terms of age, 60% fell within the 17 to 20 years range, 32.5% were aged between 20 and 25, and 7.5% were older than 25. Participants self-assessed their English proficiency, with 52.5% considering themselves at an advanced level, 37.5% at an intermediate level, and 10% identifying as beginners.

## 4 Results

Section II of the questionnaire explored students' motivations for language learning, specifically targeting instrumental, personal, and integrative motivational constructs through the first seven questions. The aim was to quantitatively assess these motivational dimensions. The statistical analysis revealed that the overall motivation of university students (N=79) to learn English is high (M=3.7), with instrumental motivation standing out as the dominant factor among the three (Table 1), although not drastically.

The highest mean values were associated with the statements related to instrumental motivation (M=3.8), followed by personal motivation, where the total mean value is also high (M=3.7). This indicates that students assess statements related to learning English



reasons for learning English, such as applying for a job or passing a test, which aligns with the highest mean score for instrumental motivation gained from the quantitative part of the research (Table 2). There were fewer responses in favour of personal motivation, with only a couple of students expressing motivations for learning English such as a genuine interest in books, writing, and gaming. Moreover, a few responses indicated support for integrative motivation, with students highlighting globalization as a key motivator. They expressed a desire to learn English to connect with people from diverse cultures globally, immerse themselves in different perspectives, and build meaningful relationships with individuals from various backgrounds.

In conclusion, the findings of this segment of the research indicate that our respondents primarily viewed English language learning as a means to enhance career prospects, with instrumental motivation scoring the highest of the three, which was confirmed by thematic analysis conducted for the open-ended question in Part II, as the students gave numerous practical reasons for learning English as major motivational drivers.

Part III of the questionnaire consists of seven questions, designed to assess students' attitudes towards ChatGPT as a writing assistance tool. Three questions are defined as yes-no questions, another three questions ask respondents to rate their ChatGPT frequency of use, preferred types of activities and usefulness of the tool's features, and one open-ended question addresses major advantages and downsides of the app.

The responses to Question 1 "How often have you used ChatGPT as a tool to assist with learning written assignments in the English language?" (Figure 1) reveal that a significant portion of students, specifically 46.3%, very frequently used ChatGPT for their assignments, while 38.8% reported not having used the app at all (or almost never). Additionally, 15% opted for "sometimes". This indicates a substantial variation in the frequency of ChatGPT utilization among the surveyed students, as the share of those who have used it and have not used it at all is rather similar, so these data suggest a rather strong polarization in the adoption of ChatGPT among the surveyed students.

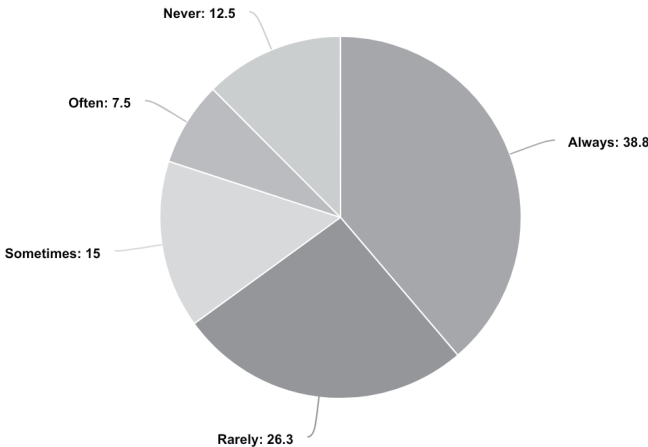


FIGURE 1. Frequency of ChatGPT use for written assignments.

Question 2 was designed to gather information on the specific tasks for which respondents most frequently utilize ChatGPT (Figure 2). The results highlight that a significant proportion of students employ the tool for various purposes, with the breakdown as follows: business emails 32%, academic purposes (essays and research papers) 27.5%, generating examples of essays 23%, other purposes 10.5%, and creative writing 7%. When the data are grouped based on similarities of areas, a prominent trend emerges, indicating that business and academic purposes are the primary areas where ChatGPT is utilized extensively (59.5%). This underscores that the majority of students deploy ChatGPT with the specific goal of achieving tasks related to both professional and educational contexts, which aligns with the earlier findings of this research, where respondents reported being motivated to learn English mainly for academic and business-related purposes.

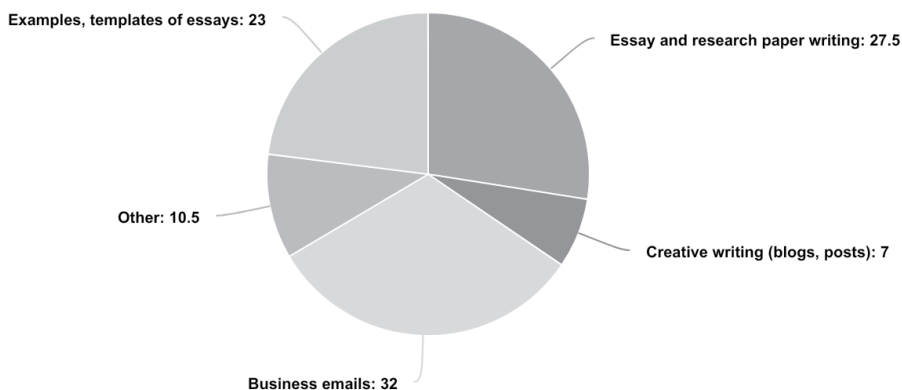


FIGURE 2. Tasks mainly performed using ChatGPT.

TABLE 3. ChatGPT for writing.

3. Evaluate ChatGPT as a tool for assistance in writing	
Valid	79
Missing	0
Mean	3.72
St. Deviation	1.250
Minimum	1.00
Maximum	5.00

The findings for question 3 (“On a scale from 1 to 5, evaluate ChatGPT as a tool for assistance in English writing”) indicate that, overall, the respondents perceive ChatGPT as a valuable resource for writing tasks ( $M=3.7$ ) (Table 3). We also sought to explore the typical applications of the tool (question 4) within the context of their written assignments for the course, and the respondents were provided with eight options that best illustrate the activity undertaken within the writing process (brainstorming, title suggestion, proofreading and sentence correction, vocabulary and word choice, summarizing/paraphrasing, text structuring,

providing arguments, and word count), which were to be rated 1–5 (Table 3). The results revealed that all the offered features scored similar means, slightly above 3, which indicates that students expressed moderate to high satisfaction with the listed features. However, the highest mean score was attributed to the “proofreading/sentence correction” option ( $M=3.5$ ), followed by “text structuring” ( $M=3.4$ ), whereas “providing arguments” scored  $M=3.2$  which emerged as the least valued functionality out of the eight that were given (Table 4).

TABLE 4. Descriptive Statistics for ChatGPT’s features.

		Brainstorming	Title suggestions	Proofreading and sentence correction	Vocabulary examples and word choice	Summarizing, paraphrasing	Text structuring	Providing arguments	Word count
N	Valid	79	79	79	79	79	79	79	79
	Missing	0	0	0	0	0	0	0	0
Mean		3.3797	3.3797	3.5190	3.4051	3.3924	3.4557	3.2025	3.3165
Std. Deviation		1.32340	1.35215	1.37618	1.29614	1.34372	1.38488	1.36231	1.40095
Minimum		1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Maximum		5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00

Question 5 was “Were all the responses provided by the tool (when given a task) reliable and accurate in terms of the written assignments you have completed?” The results indicate a noteworthy division in the respondents’ perceptions (Figure 3), as an almost equal share felt the responses provided by ChatGPT were reliable and unreliable, with 45% of the respondents stating ChatGPT was reliable, and 55% that it was not. This indicates a substantial split in opinions among the surveyed individuals regarding the trustworthiness of ChatGPT-generated content for their written assignments.

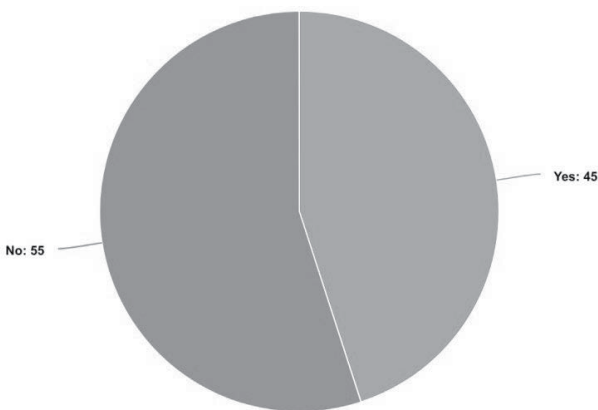


FIGURE 3. Reliability and accuracy of ChatGPT responses.

The findings from question 6, designed to evaluate respondents' willingness to recommend ChatGPT to others with writing assignments (Figure 4), present an intriguing insight. Notably, the collective percentage of those who responded "No" (8.8%) and those who expressed insecurity (28.7%) comprises almost 40% of the sample. This suggests a substantial portion of participants have reservations or uncertainties about promoting this tool to others. This hesitancy is likely influenced by the fact that nearly half of the respondents consider ChatGPT unreliable for writing assignments, as indicated in their earlier responses. These results emphasize the importance of addressing concerns related to the tool's reliability to enhance its acceptability and recommendations among users engaged in writing assignments.

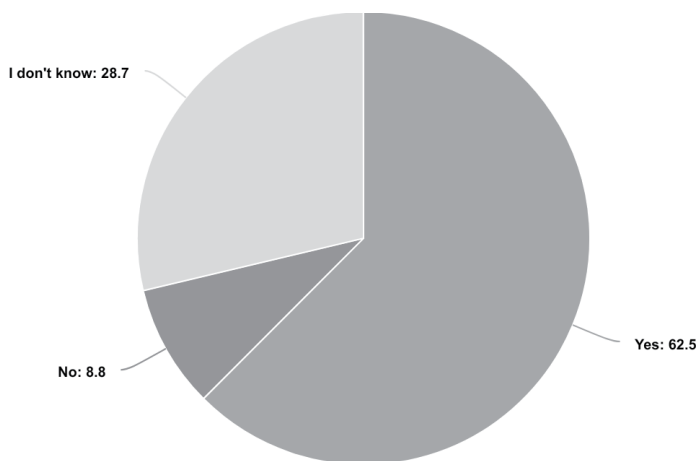


FIGURE 4. Would you recommend the tool to others?

Finally, the last question in the questionnaire was to identify the main advantages and/or disadvantages of the app. Based on the responses gathered from the questionnaire (after thematic analysis was conducted), it can be concluded that users highly appreciate the tool's efficacy and speed, accuracy in grammar, vocabulary and the information it provides, with the obvious simplicity of use (Table 5). When asked to list the most prominent downside of the tool, the most frequent response the respondents provided was "the app has no limitations", mentioned by as many as 19 of them. On the other hand, however, several "againsts" were listed, such as the inaccuracy of the data provided by the tool, the inability to tackle more complex tasks and inconsistency. The user feedback reflects a mix of positive experiences and reservations about the application's functionalities and reliability, but, overall, it seems that although a certain number of students reported that they encountered several limitations of the app, a large number of users have not yet discovered the downsides of the tool. One of the reasons for the discrepancy might also be that users have varying expectations and different language competences that affect their use of the tool. Those who primarily utilize ChatGPT for straightforward language tasks may not encounter or be bothered by its limitations, while others seeking more complex capabilities may find the tool falling short. In other words, these findings suggest that the application's effectiveness is context-dependent, and its limitations may become more apparent in specific use cases.

TABLE 5. ChatGPT advantages and disadvantages.

Advantages	Disadvantages
*Efficacy/speed – 8 responses	*No limitations – 19 responses
*Accuracy/precision of grammar/vocab/information – 7 responses	*Accuracy/incorrectness of data/inconsistency – 10 responses
*Simplicity of use – 6 responses	*I am not sure – 6 responses
*Provides a native-like communicator/speaker/responder/interactivity – 5 responses	*Repetition – 2 responses
*Know-it-all tool/knowledgeable resource – 4 responses	*Can't deal with more complex tasks – 2 responses
*Money-saver/time-saver (no tutoring) – 4 responses	*Can be used as a helpful tool only – 2 responses
*Vocabulary builder/fast and accurate translation – 4 responses	*Misses human-like quality – 1 response
*Always accessible – 2 responses	

To answer the first and the second research questions posed in this study, the researchers conducted the following additional analyses to gain a deeper understanding of the relationship between motivation, English proficiency, and the use of ChatGPT for written assignments: (1) Pearson correlation, to assess whether there is a correlation between motivation and use of ChatGPT for written assignments; (2) a one-way analysis of variance (ANOVA), to assess whether there are differences in the assessment of three motivational constructs between participants who assessed their English proficiency as basic/intermediate and advanced; and (3) a one-way analysis of variance (ANOVA), to assess whether there are differences in the assessment of ChatGPT as a writing assistance tool between participants who assessed their English proficiency as basic/intermediate and advanced. These analyses helped us explore whether students who are more motivated tend to use ChatGPT more frequently or effectively for their writing tasks, and whether there are differences in how participants perceive ChatGPT as a writing assistance tool based on their self-assessed English proficiency levels.

The results obtained using Pearson's correlation reveal (Table 6) that there is no significant correlation between students' motivation to learn English and their usage of ChatGPT for writing assignments (Table 6). This implies that the students' overall motivation levels for learning English do not necessarily predict their inclination or frequency of using ChatGPT as a writing assistance tool. This finding underscores the independence of these two aspects, indicating that factors influencing students' motivation in language learning may not directly align with their preferences or reliance on technology, such as ChatGPT, for enhancing their writing skills.

A one-way analysis of variance (ANOVA) was conducted to examine potential differences in the assessment of three motivational constructs for learning English (instrumental:  $F(1,77)=0.025$ ,  $p=0.875$ ; personal:  $F(1,77)=0.125$ ,  $p=0.724$ ; integrative:  $F(1,77)=0.255$ ,  $p=0.615$ ) as well as

TABLE 6. Correlations between motivational constructs and satisfaction with ChatGPT

		Instrumental Motivation	Personal Motivation	Integrative Motivation	Total Motivation	ChatGPT Mean
Instrumental Motivation	Pearson Correlation	1	.642**	.385**	.884**	.140
	Sig. (2-tailed)		.000	.000	.000	.220
Personal Motivation	Pearson Correlation	.642**	1	.514**	.865**	.035
	Sig. (2-tailed)	.000		.000	.000	.759
Integrative Motivation	Pearson Correlation	.385**	.514**	1	.698**	-.030
	Sig. (2-tailed)	.000	.000		.000	.790
Total Motivation	Pearson Correlation	.884**	.865**	.698**	1	.078
	Sig. (2-tailed)	.000	.000	.000		.496
DR GPT Mean	Pearson Correlation	.140	.035	-.030	.078	1
	Sig. (2-tailed)	.220	.759	.790	.496	
**. Correlation is significant at the 0.01 level (2-tailed).						

total motivation ( $F(1,77)=0.011$ ,  $p=0.915$ ) between participants who perceived their English proficiency as basic/intermediate and those who considered it advanced. Similarly, ANOVA was performed to investigate potential differences in the assessment of ChatGPT as a writing assistance tool between participants who rated their English proficiency as basic/intermediate and advanced ( $F(1,77)=0.180$ ,  $p=0.673$ ). No significant differences were observed in either analysis, indicating that the students' motivation levels and their evaluation of ChatGPT as a writing tool did not significantly vary based on their perceived English proficiency.

This suggests that students' motivation levels and their perceptions of ChatGPT's utility in enhancing writing skills remain consistent across different self-assessed proficiency levels. These findings contribute valuable insights into the robustness of students' motivations and evaluations, emphasizing the universality of ChatGPT's perceived effectiveness as a writing aid, irrespective of varying levels of English proficiency.

## 5 Discussion

### 5.1 Overview of Significant Findings

A mixed-method study was conducted with a sample of 79 IT students to investigate the relationship between motivation, proficiency in English and attitudes towards ChatGPT as

a writing assistance tool in ESP. The findings reveal that the most developed motivation is instrumental (seen as a desire to gain social recognition or economic advantages), which aligns with the attitudes towards ChatGPT's usage, underlining the tool's practical appeal. It could have been anticipated that students attending English for Specific Purposes (ESP) courses, which specifically target English for work-related purposes, would likely demonstrate a utilitarian attitude towards the utilization of ChatGPT. This finding aligns with previous research; for instance, Ardeo (2016) highlighted higher instrumental motivation among engineering students, as did Wimolmas (2013). However, contrasting results were reported by Amengual-Pizarro (2017), suggesting that integrative motivation surpasses instrumental motivation.

The findings also indicate a lack of significant correlation between overall motivation to learn English and the frequency of ChatGPT usage for writing assignments. The lack of correlation between students' motivation and their attitudes toward using ChatGPT for writing tasks could be attributed to the fact that students perceive ChatGPT as a tool for dodging responsibilities, essentially taking over the writing process on their behalf. In this scenario, the tool serves more as a convenience rather than a motivational factor, reducing the connection between motivation levels and satisfaction with ChatGPT. The awareness of its limitations may contribute to a pragmatic acknowledgement that, while handy, ChatGPT does not replace the intrinsic motivation required for genuine language learning experiences. However, it is important to note that this finding is contrary to the results obtained by Ali et al. (2023), who found correlations between motivation and ChatGPT usage in language learning (especially when it comes to reading and writing). Similarly, the research conducted by Yildiz (2023) demonstrated a connection between motivation and ChatGPT usage, which does not align with the results obtained in this study.

Furthermore, the absence of substantial differences in the assessment of motivational constructs and ChatGPT as a writing tool based on self-assessed English proficiency levels underscores the consistency of students' perceptions across varying language proficiency levels. These insights collectively emphasize the tool's potential universality, promoting its integration into language learning environments as a versatile aid that accommodates diverse motivations and proficiency levels.

Another noteworthy aspect that emerges from this research is the polarization observed in attitudes, where users express either strong positive sentiments or reservations, emphasizing the diverse perspectives on this technology. In summarizing the participants' assessments of ChatGPT, the tool's versatility and adaptability emerge as significant strengths, with users praising its perceived lack of limitations, efficacy, speed, and user-friendly interface. The accuracy and precision in grammar, vocabulary, and information contribute positively to its utility for writing assignments. However, concerns about the potential inconsistency in information and its inability to handle more complex tasks are raised as notable drawbacks. Finally, while some appreciate the open-ended nature of ChatGPT, others view it as a potential limitation.

## 5.2 Practical Implications

The practical implications of this research imply that ChatGPT can be a useful writing tool for students across different English proficiency levels. Teachers and students can thus consider

incorporating it into writing tasks without concerns about varying proficiency levels affecting its perceived usefulness. Secondly, students' motivation to learn English remains important, regardless of their perceived proficiency, so educators should focus on fostering motivation as a general factor for language learning success, rather than tailoring strategies based on proficiency levels. Given that proficiency levels did not significantly impact the perception of ChatGPT as a writing aid, lecturers can encourage students to utilize such tools for writing support, promoting inclusivity and accessibility in language learning environments, while remaining attentive to individual preferences and needs. Some students might still benefit more from personalized assistance or alternative writing tools based on their learning styles. In essence, the study suggests that tools like ChatGPT can be widely useful for students in improving their writing skills, regardless of how well they think they know English. It also emphasizes the importance of motivation and opens up possibilities for more inclusive and accessible language learning environments. Finally, the fact that the respondents confessed to having an instrumental motivation for learning English highlights the fact that language instructors should leverage the instrumental motivation for writing tasks and emphasize how using such tools aligns with their practical goals, such as achieving better academic results or preparing for future careers. Finally, as strong polarization is seen in several responses related to the limitations of the reliability of the tool, it is essential that informed guidance on how regular use can lead to tangible improvements in line with the students' practical aspirations be provided.

## 6 Limitations and Future Study

The study relies on self-reported data, which can be subject to biases and inaccuracies. As the researchers did not witness the actual use of the tool in written assignments, the participants may have provided responses based on perception rather than objective experiences, influencing the validity of the results. Moreover, the study's results are contingent on the characteristics of the sample population, so if the participants are not representative of a broader demographic, the generalizability of the findings may be limited. Finally, the study's findings may be time-sensitive as technology evolves rapidly, so ChatGPT's features constantly improve and provide a better user experience.

Building on the current study, future research could explore several avenues to further enhance our understanding of the dynamics between language learning, motivation, and the use of writing assistance tools: comparing the effectiveness and user experiences of ChatGPT with other writing assistance tools or language models could offer a more comprehensive understanding of the relative advantages and limitations of different technologies. Moreover, investigating the impact of targeted user training on the effectiveness and user satisfaction with writing assistance tools could also be beneficial. Additionally, intervention studies could explore how specific interventions influence motivation and language learning outcomes.

Finally, further studies on ethical considerations would additionally emphasize the crucial need for responsible implementation and guidance when integrating ChatGPT into writing assignments. It is imperative to ensure that students are well-informed about the tool's limitations and ethical implications, preventing overreliance and fostering a balanced approach to incorporating this technology into their writing tasks.

## 7 Conclusion

The study aimed to explore the interplay between motivation for English language learning and attitudes towards ChatGPT as a writing assistance tool in the context of English for Specific Purposes (ESP) among IT students. Through a mixed-method approach involving 79 IT students the research sought answers to three pivotal questions: identifying prevalent types of motivation among students and their connection with ChatGPT usage frequency, exploring how English proficiency levels shape attitudes towards ChatGPT, and capturing students' perspectives on ChatGPT within the framework of completing homework writing assignments. The research instrument combined adapted motivation scales and original questions. Data were collected via Google Forms and subsequently analysed using a combination of statistical techniques, including SPSS for descriptive statistics, ANOVA for examining differences between groups, Pearson correlation coefficient analysis for assessing relationships between variables, and thematic analysis techniques for extracting meaningful insights from qualitative responses.

The analysis revealed instrumental motivation as predominant, aligning with practical attitudes towards ChatGPT's usage. However, a lack of significant correlation between motivation for learning English and ChatGPT usage frequency suggested that students may perceive the tool as a convenience rather than a motivational factor. Moreover, it was found that students' motivation levels and their assessment of ChatGPT as a writing tool did not exhibit significant variance in relation to their perceived English proficiency. This highlighted ChatGPT's potential universality, accommodating diverse learners, notwithstanding the polarization observed in user attitudes. Additionally, the research revealed that students recognized several strengths of ChatGPT, such as its versatility, speed, and user-friendliness, while simultaneously acknowledging concerns regarding information consistency and handling complexity. These findings contribute to understanding ChatGPT's role as a writing aid and highlight areas for further exploration in integrating such technologies into language learning environments, which is now increasingly recognized as indispensable.

Without any doubt, the emergence of ChatGPT, a groundbreaking language model and a demonstration of the potential of AI, has significantly infiltrated all spheres of life, evoking a spectrum of opinions. The tool has sparked contrasting views within academic settings, with students mainly expressing excitement while educators have been raising ethical concerns about its potential misuse, and how it might even break the current educational system. One thing is certain – the landscape of writing tasks within language learning has been altered forever. The discourse surrounding this topic is multifaceted: some voice reservations about the potential obsolescence of traditional writing assignments, while others perceive ChatGPT as a transformative catalyst capable of helping users overcome writer's block and thus enhancing writing proficiency.

As educators and developers move forward, acknowledging these varied attitudes and addressing concerns could be pivotal in fostering a more inclusive and effective integration of AI-driven writing tools in language learning environments. As language instructors caught in the moment of AI technology infiltrating so many educational segments, we believe that a continuous dialogue between stakeholders and a commitment to user-centred design will be

essential in shaping the future landscape of language education technology, ensuring that it aligns with the diverse needs and expectations of students in the evolving digital era.

## References

- Ali, Jamal Kaid Mohammed, Muayad Abdulhalim Ahmad Shamsan, Taha Ahmed Hezam, and Ahmed A. Q. Mohammed. 2023. "Impact of ChatGPT on learning motivation: Teachers and students' voices." *Journal of English Studies in Arabia Felix* 2 (1): 41–49. <https://doi.org/10.56540/jesaf.v2i1.51>.
- Al-Tamimi, Atef, and Munir Shiub. 2009. "Motivation and attitudes towards learning English: A study of petroleum engineering undergraduates at Hadhramout University of Sciences and Technology." *GEMA Online Journal of Language Studies* 9 (2): 29–41.
- Amengual-Pizarro, Marian. 2017. "Engineering students' motivational variables towards English and the learning of the English language." *LFE - Revista De Linguas Para Fines Especificos* 23 (1): 31–44.
- Ardeo, José Miguel González. 2016. "Engineering students' instrumental motivation and positive attitude towards learning English in a trilingual tertiary setting." *Ibérica* 32: 179–200.
- Athanassopoulos, Stavros, Polyxeni Manoli, Maria Gouvi, Konstantinos Lavidas, and Vassilis Komis. 2023. "The use of ChatGPT as a learning tool to improve foreign language writing in a multilingual and multicultural classroom." *Advances in Mobile Learning Educational Research* 3 (2): 818–24. <https://doi.org/10.25082/AMLER.2023.02.009>.
- Baskara, Risang. 2023. "Exploring the implications of ChatGPT for language learning in higher education." *Indonesian Journal of English Language Teaching and Applied Linguistics* 7 (2): 343–58.
- Brown, H. Douglas. 2000. *Principles of Language Learning and Teaching*. 4th ed. White Plains: Longman.
- Byrd, Christine. 2023. "WRITE center helps teachers grapple with ChatGPT and future of writing instruction." *UCI School of Education*, February 6, 2023. <https://education.uci.edu/chatgpt.html>.
- Byrne, Donn. 1988. *Teaching Writing Skills*. London: Longman Press.
- Cahyono, Bambang Yudi, and Titik Rahayu. 2020. "EFL students' motivation in writing, writing proficiency, and gender." *TEFLIN Journal: A Publication on the Teaching and Learning of English* 31 (2): 162–80. <https://doi.org/10.15639/teflinjournal.v31i2/162-180>.
- Chen, Wenting. 2021. "Understanding students' motivation in L2 collaborative writing." *ELT Journal* 75 (4): 442–50. <https://doi.org/10.1093/elt/ccab027>.
- Claybourn, Cole. 2023. "Why some college professors are embracing ChatGPT." *U. S. News*, May 9, 2023. <https://www.usnews.com/education/best-colleges/articles/why-some-college-professors-are-embracing-chatgpt>.
- Cooper, L. Robert, and Joshua A. Fishman. 1977. "A study of language attitudes." In *The Spread of English: The Sociology of English as an Additional Language*, edited by J. A. Fishman, R. L. Cooper, and A. W. Conrad, 239–76. Rowley: Newbury House.
- Crookes, Graham, and Richard W. Schmidt. 1991. "Motivation: Reopening the research agenda." *Language Learning* 41 (4): 469–512.
- Deci, Edward L., and Richard M. Ryan. 1985. *Intrinsic Motivation and Self-Determination in Human Behaviour*. New York: Plenum.
- Dörnyei, Zoltan. 1990. "Conceptualizing motivation in foreign-language learning." *Language Learning* 40 (1): 45–78.
- . 1994. "Motivation and motivating in the foreign language classroom." *The Modern Language Journal* 78 (3): 273–84.
- Ellis, Rod. 2015. *Understanding Second Language Acquisition*, 2nd Edition. Oxford: Oxford University Press.
- Fareed, Muhammad, Almas Ashraf, and Muhammad Bilal. 2016. "ESL learners' writing skills: Problems, factors and suggestions." *Journal of education and social sciences* 4 (2): 81–92.
- Fitria, Tira Nur. 2023. "Artificial intelligence (AI) technology in OpenAI ChatGPT application: A review of ChatGPT in writing English essay." *ELT Forum: Journal of English Language Teaching* 12 (1).
- Gardner, Robert C., and Wallace E. Lambert. 1972. *Attitudes and Motivation in Second Language Learning*. Rowley, MA: Newbury House Publishers.

- Gardner, Robert C. 1983. "Learning another language: A true social psychological experiment." *Journal of Language and Social Psychology* 2: 219–40.
- . 1985. *Social Psychology and Second Language Learning: The Role of Attitude and Motivation*. London: Edward Arnold.
- Hardianti, Meisela, and Banatul Murtafi'ah. 2022. "Unveiling attitude and motivation of English Education students toward English." *Journal of English and Education* 8 (4): 37–46.
- Harmer, Jeremy. 1998. *The Practice of English Language Teaching*. London: Pearson.
- Hedge, Tricia. 1988. *Writing*. Oxford: Oxford University Press.
- Hyland, Ken. 2003. "Genre-based pedagogies: A social response to process." *Journal of Second Language Writing* 12 (1): 17–29. [https://doi.org/10.1016/S1060-3743\(02\)00124-8](https://doi.org/10.1016/S1060-3743(02)00124-8).
- Kohnke, Lucas, Benjamin Luke Moorhouse, and Di Zou. 2023. "ChatGPT for language teaching and learning." *RELC Journal* 54 (2): 537–50. <https://doi.org/10.1177/00336882231162868>.
- Kostka, Ilka, and Rachel Toncelli. 2023. "Exploring Applications of ChatGPT to English language teaching: Opportunities, challenges, and recommendations." *TESL-EJ* 27 (3). <https://doi.org/10.55593/ej.27107int>.
- Ling, Yin Ling Cheung. 2016. "Teaching Writing." In *English Language Teaching Today: Building a Closer Link between Theory and Practice*, edited by W. A. Renandya, and H. P. Widodo, 1–20. New York: Springer International.
- Marche, Stephen. 2022. "The college essay is dead." *The Atlantic*, December 6, 2022. <https://www.theatlantic.com/technology/archive/2022/12/chatgpt-ai-writing-college-student-essays/672371/>.
- McDonough, Steven. 1981. *Psychology in Foreign Language Teaching*. London: George Allen & Unwin.
- Meniado, Joel C. 2023. "The impact of ChatGPT on English Language teaching, learning, and assessment: A rapid review of literature." *Arab World English Journals (AWEJ)* 14 (4). <https://doi.org/10.24093/awej/vol14no4.1>.
- Nugroho, Arif, Nur Hidayanto Pancoro Setyo Putro, and Kastam Syamsi. 2023. "The potentials of ChatGPT for language learning: Unpacking its benefits and limitations." *REGISTER JOURNAL* 16 (2): 224–47. <https://doi.org/10.18326/register.v16i2.224-247>.
- Oliveira, Luciana, and Tony Silva. 2016. "Second language writing in elementary classrooms: An overview of issues." In *Second Language Writing in Elementary Classrooms*, edited by Luciana Oliveira, and Tony Silva, 1–10. London: Palgrave Macmillan.
- Perez, Julianna. 2023. "ChatGPT is destroying education." *Eagle Media*, November 20, 2023. <https://eaglenews.org/30720/opinion/chatgpt-is-destroying-education/>.
- Raimes, Ann. 1983. *Techniques in Teaching Writing*. Oxford: Oxford University Press.
- Roose, Kevin. 2023. "Don't ban ChatGPT in schools. Teach with it." *The New York Times*, January 12, 2023. <https://www.nytimes.com/2023/01/12/technology/chatgpt-schools-teachers.html>.
- Smith, Robert Reed. 2023. "Harness the apocalyptic power of AI for your fiction." *Dabble*, April 20, 2023. <https://www.dabblewriter.com/articles/chatgpt-for-fiction>.
- Spatford, Sam. 2023. "ChatGPT's apocalypse already happened." *The Amherst Student*, February 15, 2023. <https://amherststudent.com/article/chatgpts-apocalypse-already-happened/>.
- Tica, Lena, Ivana Krsmanović, and Jenny Lemper. 2023. "ChatGPT as a double-edged sword or a game-changer: Revolution in ESL classroom." A discussion at the *Peralta Equity conference*, 26–28 April, 2023. <https://www.youtube.com/watch?v=SgNGukT6fyk>.
- Ulla, Mark Bedoya, William F. Perales, and Stephenie Ong Busbus. 2023. "'To generate or stop generating response': Exploring EFL teachers' perspectives on ChatGPT in English language teaching in Thailand." *Learning: Research and Practice* 9 (2): 168–82. <https://doi.org/10.1080/23735082.2023.2257252>.
- Vukićević-Đorđević, Ljiljana. 2015. "Emerging technologies: Does it feel like learning?" *The Journal of Teaching English for Specific and Academic Purposes* 3 (3): 483–97.

- Wimolmas, Ratanawalee. 2013. *A Survey Study of Motivation in English Language Learning of First Year Undergraduate Students at Sirindhorn International Institute of Technology (SIIT), Thammasat University*. Thailand: Language Institute, Thammasat University.
- Xiao, Yangyu, and Yuying Zhi. 2023. "An exploratory study of EFL learners' use of ChatGPT for language learning tasks: Experience and perceptions." *Languages* 8 (3): 212. <https://doi.org/10.3390/languages8030212>.
- Yıldız, Tugba Aydın. 2023. "The impact of ChatGPT on language learners' motivation." *Journal of Teacher Education and Lifelong Learning* 5 (2): 582–97. <https://doi.org/10.51535/tell.1314355>.
- Zadorozhnyy, Artem, and Wan Yee Winsy Lai. 2024. "ChatGPT and L2 written communication: A game-changer or just another tool?" *Languages* 9 (1): 5. <https://doi.org/10.3390/languages9010005>.

## Appendix: The Questionnaire

No.	<b>I section: Demography</b>	
1.	How old are you?	
2.	What is your gender?	
3.	Rate your proficiency in the English language.	
	<b>II section: Motivation for learning English</b>	
1.	Instrumental motivation	I learn English because it will enable me to carry my tasks more efficiently
2.		I learn English because I hope to further my education
3.		I learn English because it is a university requirement
4.	Personal motivation	I learn English for a personal development
5.		I learn English because it will enhance my status among my friends
6.	Integrative motivation	I learn English in order to integrate with the Western culture
7.	I learn English for some other reason. What? Please specify.	
	<b>III section: ChatGPT</b>	
1.	How often have you used ChatGPT as a tool to assist with learning written assignments in the English language? a) always b) rarely c) sometimes d) often e) never	
2.	Which tasks do you most frequently use ChatGPT for?	
	- For academic essays, for writing seminar/research papers	
	- For creative writing (blogs, posts, essays, personal emails, etc.)	
	- For professional communication (business emails)	
	- For brainstorming ideas, generating essay examples	
	- For other purposes	

3.	On a scale from 1 to 5, evaluate ChatGPT as a tool for assistance in English writing.
4.	On a scale from 1 to 5, rate the usefulness of the following segments of this tool:
	- Brainstorming and generating ideas for writing.
	- Title suggestions
	- Proofreading and sentence correction
	- Vocabulary examples and word choice
	- Summarizing, paraphrasing
	- Text structuring
	- Providing arguments
	- Word count
5.	Were all the responses provided by the tool (when given a task) reliable and accurate, in terms of the written tasks you have completed?
6.	Would you recommend this tool to others as assistance for writing tasks?
7.	What are the advantages and disadvantages of this tool, based on your experience? Please specify.



## Investigating Oral Presentation Assessment in LSP: Practices, Demands and Challenges

### ABSTRACT

Giving an oral presentation is a widely used assessment task in higher education, and is very efficient in preparing LSP students for future professional contexts. Empirical studies have made significant contributions to revealing students' attitudes towards the benefits of oral presentations, but very few studies have investigated LSP teachers' attitudes regarding the assessment of such presentations. This paper aims to provide deeper insights into LSP teachers' assessment practices, needs, demands and challenges when evaluating oral presentations. For that purpose, 103 LSP teachers in Serbia and Croatia were asked to indicate the greatest challenges in the assessment process and the most suitable solutions for improving evaluations. The results show that the teachers have difficulty defining the significance level of each rating criterion that describes a student's performance. In line with this, developing a unique rating scale specifically for LSP oral presentation assessment is recognized as the most helpful strategy in dealing with the challenges they face.

**Keywords:** oral presentation, assessment, LSP, higher education context

## Ocenjevanje govornih nastopov pri poučevanju jezika stroke: izkušnje, zahteve in izzivi

### IZVLEČEK

Ocenjevanje govornih nastopov predstavlja sestavni del poučevanja jezika stroke (LSP) v visokem šolstvu, za študente pa predstavlja pomemben del priprav na poklicno kariero. Odnos študentov do govornih nastopov obravnavajo številne empirične raziskave, le redke pa stališča učiteljev jezika stroke o ocenjevanju govornih nastopov. Namen pričujočega prispevka je prikazati izkušnje učiteljev z ocenjevanjem govornih nastopov, njihove potrebe, pričakovanja in izzive pri ocenjevanju govornih nastopov. V raziskavi so sodelovali 103 učitelji jezika stroke v Srbiji in na Hrvaškem in izpostavili največje težave pri ocenjevanju ter predlagali rešitve za izboljšanje ocenjevanja govornih nastopov. Rezultati so pokazali, da imajo učitelji težave z določanjem stopnje pomembnosti vseh posameznih kriterijev ocenjevanja govornih nastopov. Učitelji menijo, da je najboljša strategija za ocenjevanje govornih nastopov priprava enotne ocenjevalne lestvice.

**Ključne besede:** govorni nastop, ocenjevanje, jezik stroke (LSP), visoko šolstvo

# 1 Introduction

Public speaking has a long teaching tradition, and has been studied and practised since ancient Greek times, and today it serves as the foundation for many undergraduate curricula focusing on speaking and communication (Vangelisti, Daly, and Friedrich 1999, 75–80). Since public speaking is a common pedagogical task, some practical issues seem to appear as constant challenges: making a choice of suitable approaches and teaching methods, developing precise rating criteria, encouraging students to participate in speaking activities, creating appropriate study courses, and fostering a positive learning environment. In contemporary society, oral presentation represents a core public speaking activity undertaken at the tertiary level of education (Girard, Pinar, and Trapp 2011; Alwi and Sidhu 2013; Alshammari and Mugaddam 2023). In an ever-changing educational context, oral presentation remains one of the most widely used tasks even in technologically advanced contexts, and research indicates that both teachers and students benefit substantially from oral presentation tasks within an online setting (Hill and Storey 2003, Petrović, Radović Firat, and Palurović 2022). However, it is necessary to provide more evidence of teachers' attitudes about oral presentation teaching and assessment practices (Alshammari and Mugaddam 2023, 133). This would help instructors and syllabi creators gain a deeper understanding of the methodology necessary for designing adequate teaching materials, and teaching and assessment approaches.

Several studies highlight the benefits of students' self-assessment of ESP oral presentations in higher education, such as raising students' awareness of their own and their peers' proficiency (Rizvi 2004), and emphasize the need for both students and teachers to focus on areas of oral performance that require further improvement (Burkšaitienė 2020). Furthermore, Palmour (2024) investigates English for Academic Purposes (EAP) teachers' practices when assessing university students' presentations. The author considers that basic construct-irrelevant features, which sometimes guide the assessment process, can be addressed by using rating scales and providing teacher training in the context of EAP assessment, leading to better decision-making by teachers. On top of this, we wanted to gain a broader overview of teachers' practices, needs, demands and challenges when evaluating oral presentations in a language for specific purposes (LSP) context, since needs analysis is one of the most prominent characteristics of LSP (Basturkmen and Elder 2004, 674). We thus considered that a survey conducted with 103 LSP teachers in the region of Serbia and Croatia would help to reveal specific requirements of the oral presentation assessment process.

## 2 Literature Review

### 2.1 Definition and Characteristics of an Oral Presentation Task

An oral presentation is an open-ended task (Luoma 2004) that enables learners to express their ideas, beliefs and attitudes to members of an audience. According to the Common European Framework of Reference for Languages (CEFR), oral production/speaking activity represents a speaker's production of spoken content for an audience (Council of Europe 2001, 58). Speaking activities encompass public addresses, speeches at public events, university lectures, sermons, entertainment, sports commentaries, presentations, etc. Learners can read a text, speak from notes, use visual aids to support their speech, act or sing. Similarly,

Richards (2008, 27) defines oral presentation as a spoken activity performed in front of an audience. Polovina and Dinić (2014, 410) also state that an oral presentation is a form of public performance and explain that numerous activities occur before the final delivery of the speech. More precisely, in order to perform a presentation, learners engage in wide reading, writing, and translation if necessary. The delivery of a speech is interactive. Clearly, oral presentation is a complex and demanding task that includes receptive, productive and interactive activities, as well as preparation and practising strategies.

To prepare and perform a presentation, learners gradually scaffold the necessary skills and strategies (Morita 2000, 294). Polovina and Dinić (2014, 410) list the following stages of oral performance preparation: 1) choosing a topic, 2) data collection, 3) reflecting on the acquired information, 4) selecting topic-related information, 5) creating a presentation, and 6) preparation for speech delivery. The final phase refers to public speaking, or oral presentation followed by the teacher's evaluation, peer assessment or self-assessment.

Richards (2008, 28) discerns the following features of oral presentations: 1) addressing an audience, 2) creating a positive impression with the audience and engaging the audience to interact, 3) following logical structure and organization of presentation with a suitable introduction and conclusion, 4) using a monologue as a specific linguistic format, 5) employing appropriate style and register, 6) accuracy of pronunciation and grammar, and 7) selecting the right vocabulary suitable for the context or audience. This leads us to the conclusion that the task of oral presentation encompasses the language parameters, the quality of speaker-audience interaction and the structure and organization of the presentation. Hence, the evaluation of a student's performance cannot be conducted through traditional testing, but rather through performance-based assessment.

## 2.2 Oral Performance Assessment

Norris et al. (1998, 12) explain that the assessment of any task requiring language use and performance simultaneously is considered a performance-based assessment. Assessment is one of the most complex and sensitive assignments, since the evaluation outcomes affect students' future educational and professional prospects. That being the case, language teaching assessment (LTA) literacy has been developed as a concept that refers to the teacher's proficiency in creating, designing, and assessing tests and various evaluation methods. It also encompasses the ability to analyse and rank scores based on theoretical principles (Vogt and Tsagari 2014, 377). Atay and Mede (2017, 44) agree with this description, adding that teachers need to acquire knowledge about various evaluation approaches and methods, assessment outcomes and consequences, as well as become familiar with both traditional and innovative assessment methods. These procedures need to be adequately adapted to suit the teaching needs. Fulcher (2012) further argues that it is vital for teachers to gain the skills needed for practical implementation of such knowledge.

However, the empirical research conducted in different parts of the world indicates that teachers' assessment practices show considerable inconsistencies (Bøhn 2015). Bøhn (2015) argues that although the teachers in his study were evaluating oral performance without a specific rating scale, they created their own rating criteria to use in the assessment processes.

Although the teachers paid attention to similar constructs, a problem appeared when they needed to evaluate the criteria concerning the level of significance attributed to each particular criterion. This further led to some discrepancies with regard to the results of scoring.

In line with this, Yildiz (2011) conducted research among Norwegian teachers and found there was little inter-rater reliability since different teachers paid attention to different aspect of the students' performance. They also evaluated the same criteria differently and included construct-irrelevant features in their assessment procedures. The same problem was confirmed by Smir's (2020) study conducted with teachers in Sweden. They found it difficult to explain the salient rating criteria used to assess the oral proficiency of their students.

Brahim, and Fadhila (2021) report several problems that FL speaking instructors come across in the classroom context, such as a demanding syllabus which does not enable teachers to engage learners in speaking activities as much as necessary, as well as coursebooks that do not promote sufficient speaking tasks and activities.

Based on the literature review provided, it is clear that evaluating oral presentations is quite complex and challenging. Teachers, as raters, encounter various requirements when assessing oral performance, with the most prominent challenge being the complexity of determining the level of significance of each specific criterion used in the assessment procedure.

### **3 Methodology**

#### **3.1 Aim of the Research**

The topic of this research pertains to teachers' practices considering the assessment of LSP oral presentations, since LSP is frequently used in academic and professional contexts (Basturkmen and Elder 2004, 672–73). Study courses in LSP at universities are organized to fulfil the requirements of specific learning needs that engage students in spoken language activities, such as seminars, presentations and various forms of interactive collaboration. In this research, we aimed to investigate the real needs, demands and challenges of the teachers, as raters, in higher education settings. In line with the theoretical grounds in the previous section, we created the following research questions:

RQ1: What is identified as the primary challenge in assessing the LSP oral presentation?

RQ2: What strategies can facilitate teachers in improving the LSP oral presentation assessment process?

#### **3.2 Instrument**

A Google Forms survey was used to collect the data for the research analysis. The survey comprised four sections: 1) demographic questions, 2) the teachers' practices when assessing LSP oral presentations, 3) the teachers' needs and challenges when assessing LSP oral presentations, and 4) the teachers' suggestions for enhancing the oral presentation assessment. The data collection period included the end of 2022 and the beginning of 2023.

The first part of the survey contains questions related to the respondents' gender, age, work experience and the language for specific purposes they teach. The second section consists of nine Likert-scale questions based on the theoretical grounds discussed in section 2 of this paper, which presented the problems that teachers in different parts of the world encountered while rating oral performances. Hence, we wanted to check whether the teachers from the region of Serbia and Croatia faced similar demands and challenges in their assessment procedures. Therefore, the first step was to find out how many raters took part in the assessment process when evaluating oral presentations in the classroom context, or whether the teachers asked for an expert's opinion on the subject content when assessing LSP presentations. We were also curious about the professional development workshops and seminars they attended related to different aspects of assessment. It was also important to determine which of the two assessment approaches, formative or summative, prevailed in the evaluation procedure, as well as whether the teachers provided corrective feedback to students in the final phase of evaluation. Additionally, the teachers were asked whether they conducted real-time assessment or whether they asked students for consent to record their performances for later assessment. The third part of the survey also included nine Likert-scale questions that investigated whether the process of the LSP oral presentation assessment was enjoyable, challenging, demanding or tiresome for the teachers. We were also interested in whether they used standardized scales, such as the CEFR rating scales or the IELTS rating scales, or if they created their assessment scales to suit the specific assessment context. The fourth part of the survey contained only one open-ended question. Here, we aimed to gain insight into the teachers' perspectives on potential mechanisms for further improvement of the oral presentation assessment process in the LSP classroom.

### 3.3 Research Respondents

The participants in this research were 103 foreign language teachers who teach LSP in higher education institutions and secondary vocational schools in the region of Serbia and Croatia. When we analyse the data obtained by the survey, it shows that most of the teachers (95.1%) are female, with only 4.9% of the respondents being male FL teachers. The greatest number of teachers are between 35 and 45 years of age (39.8%). A slightly smaller number of respondents are teachers between 45 and 55 years of age (35.9%), while there are 20.4% of teachers between 55 and 65. The number of the youngest and oldest teachers is approximately the same, i.e., 2.9% of teachers are between 25 and 35, while the teachers over 65 make up only 1% of the sample.

According to the survey, the greatest number of teachers have more than 20 years of teaching experience (55.3%), followed closely by those who have been working for more than 10 years and less than 20 (35.9%). Just 8.8% of teachers have less than 10 years of work experience.

Finally, with regard to the particular LSP field, most of the respondents teach English for Specific Purposes (57.28%), followed by English and German for Specific Purposes (11.65%), German for Specific Purposes (7.7%), Business English (2.97%), French for Specific Purposes (2.91%) and Italian for Specific Purposes (1.94%). Just one of the respondents (0.97%) teaches each of the following subjects: Technical English, English in Tourism, English in Agriculture, Technical German, Business German in Engineering, English in Aircraft, English

in Maritime Engineering, Russian for Specific Purposes, English in IT, English in Fashion Design, English in Graphic Technology, English for Managers, English in Law, German in Law, English in Medicine, and Vocational English.

## 4 Findings of the Study and Discussion

The quantitative analysis of the answers from the second part of the survey helped us to identify typical activities and approaches the teachers apply when assessing oral presentations. The respondents were asked to estimate how often they performed certain rating activities. The results are presented in Table 1.

The analysis of the results provided in Table 1 shows that most teachers always or almost always assess the students' LSP oral performances on their own (75.72%). A small number of teachers (16.50%) sometimes assess oral presentations independently, while 7.77 % of the respondents rarely or never assess presentations by themselves.

To support the previous results, we designed the second question to find out how often the teachers assessed oral presentations in collaboration with another evaluator. The results we gained are in line with the previous ones. Namely, it turns out that most teachers (83.49%) either never ask for the collaboration of another colleague, or do so rarely (9.71%). As such, a very small percentage of teachers frequently have the support of another rater in the oral presentation assessment (6.79%).

The third question was used to check how often the teachers, as raters, collaborated with an expert who was familiar with the focal topic in the presentation. We considered that this information would prove highly beneficial, as LSP teachers also need the specific field knowledge in addition to proficiency in English. These findings also show that the teachers never or rarely include a specific field expert in their rating processes (85.43%), while 11.65% of them sometimes collaborate with an expert from a specific area. Only 2.91% of raters sometimes require an expert's collaboration in evaluation. Finally, there were no teachers who always included a specific field expert in their rating processes.

Further, we wanted to determine how frequently the teachers attended training sessions and professional development programmes to improve their assessment competencies. The results proved that the teachers predominantly did not attend teacher training or professional development activities that dealt with assessment and evaluation (67.96%). Only a small number of teachers often or always attend such seminars (8.74%). Just under a quarter of the teachers, 23.30%, sometimes work to improve their assessment skills by taking part in teacher training.

The following three questions were designed to investigate the frequency of summative assessment, and certain aspects of formative assessment, in the domain of public speaking performance. The fifth question gathered data on the frequency of summative assessment. Most teachers (74.76%) stated that the grades they gave were always or often numerical, while 10.68% of respondents evaluated their students' performances numerically. Only 14.57 % of teachers never or rarely conduct summative evaluation.

TABLE 1. Teachers' practices for LSP oral presentation assessment.

Question	Responses	Number of teachers' responses	Percentage
1. How frequently do you assess oral presentations on your own?	Never	1	0.97%
	Rarely	7	6.80%
	Sometimes	17	16.50%
	Often	27	26.21%
	Always	51	49.51%
2. How frequently do you assess oral presentations in collaboration with another evaluator?	Never	62	60.19%
	Rarely	24	23.30%
	Sometimes	10	9.71%
	Often	5	4.85%
	Always	2	1.94%
3. How often do you collaborate with an expert familiar with the focal topic during the assessment of oral presentations?	Never	61	59.22%
	Rarely	27	26.21%
	Sometimes	12	11.65%
	Often	3	2.91%
	Always	0	0.00%
4. How often do you attend organized training and professional development programs to enhance competencies for assessing oral presentations?	Never	43	41.75%
	Rarely	27	26.21%
	Sometimes	24	23.30%
	Often	7	6.80%
	Always	2	1.94%
5. How often do you conduct summative assessment of oral presentations?	Never	8	7.77%
	Rarely	7	6.80%
	Sometimes	11	10.68%
	Often	20	19.42%
	Always	57	55.34%
6. In the case of summative assessment, how often do you provide feedback about your final grade?	Never	7	6.80%
	Rarely	1	0.97%
	Sometimes	9	8.74%
	Often	20	19.42%
	Always	66	64.08%
7. In the case of summative assessment, how often does your feedback contain instructions for further improvement of speech delivery?	Never	4	3.38%
	Rarely	2	1.94%
	Sometimes	7	6.80%
	Often	23	22.33%
	Always	67	65.05%
8. How often do you conduct the oral presentation assessment immediately after students have finished oral presentations?	Never	3	2.91%
	Rarely	4	3.88%
	Sometimes	9	8.74%
	Often	24	23.30%
	Always	63	61.17%
9. How often do you record oral presentations to evaluate them later?	Never	70	67.96%
	Rarely	16	15.53%
	Sometimes	7	6.80%
	Often	9	8.74%
	Always	1	0.97%

The sixth question is an extension of the previous one. Here, we intended to check whether the teachers who performed their assessment summatively also provided feedback about their decisions. The results were encouraging, since 83.50% of the raters always or often give additional explanations to students about their grades. Only 8.74% of teachers sometimes provide feedback about their students' scores. The percentage of those who never or rarely explain the final grades they give is quite insignificant (7.77%).

The subsequent inquiry aimed to establish how frequently the teachers provided a numerical grade along with feedback which would tell the students how they could improve their future oral performances. The responses show that feedback is always or often provided by 87.38% of the raters, while 6.80 % only sometimes offer feedback. The percentage of those who never or rarely instruct their students on how to improve their oral performances is quite small, at 5.82%.

Further survey analysis enabled deeper insight into the assessment time frame. More precisely, we wanted to inquire whether the teachers do a real-time assessment, simultaneously with the students' speech delivery. Another option is to request consent from the students to be recorded, allowing their presentations to be evaluated sometime later. It turns out that a great majority of teachers conduct a real-time assessment (84.47%), whereas only 8.74% of raters sometimes evaluate the students' oral performances while they are speaking. The least number of teachers rarely or never assess their learners while they are speaking. These results are confirmed by the following results, which show that 83.49% of teachers never or rarely record presentations with the purpose of subsequent assessment, while only a small percentage of raters sometimes record presentations for the sake of subsequent evaluation (6.80%). Just under a tenth of the teachers (9.71%) always or often record students' oral presentations and evaluate them later.

To sum up, the previously mentioned inferences indicate that teaching practices related to assessing oral presentations in higher education institutions in Serbia and Croatia show consistent results. Firstly, most teachers evaluate LSP oral presentations independently, without collaboration with another evaluator, or a specific field expert. Secondly, the teachers rarely attend teaching training and professional education and development sessions related to assessment. Furthermore, the prevailing approach to assessment is summative. It is complemented with formative assessment elements, which provides detailed feedback along with instructions for further improvement of oral performance. Finally, assessment is most frequently performed in the classroom, at the same time as the students are speaking.

The third segment of the survey investigated the needs and challenges faced by the raters during the oral presentation assessment process. The results of the quantitative data analysis are presented in Table 2.

The analysis of the first statement proves that most of the teachers (83.49%) completely or almost agree that they do not have any difficulties in assessing oral presentations. Only a small number of respondents (9.71%) declare that they completely or mostly agree that they have difficulties when rating their students' performances. Only 6.80 % of teachers were undecided on this matter.

TABLE 2. Needs and challenges in LSP oral presentations assessment.

Question: To what extent do the following statements refer to your assessment process?	Responses	Number of teachers' responses	Percentage
1. I have no difficulties in evaluating oral presentations.	I completely disagree.	2	1.94%
	I mostly disagree.	8	7.77%
	I am neutral.	7	6.80%
	I mostly agree.	59	57.28%
	I completely agree.	27	26.21%
2. Oral presentation assessment is interesting and quite challenging.	I completely disagree.	2	1.94%
	I mostly disagree.	13	12.62%
	I am neutral.	15	14.56%
	I mostly agree.	56	54.37%
	I completely agree.	17	16.50%
3. Oral presentation assessment is a time-consuming process and I sometimes find myself feeling fatigued.	I completely disagree.	21	20.39%
	I mostly disagree.	42	40.78%
	I am neutral.	17	16.50%
	I mostly agree.	18	17.48%
	I completely agree.	5	4.85%
4. I am undecided about whether to evaluate different presentation features equally when making inferences about the final grade.	I completely disagree.	19	18.45%
	I mostly disagree.	36	34.95%
	I am neutral.	25	24.27%
	I mostly agree.	18	17.48%
	I completely agree.	5	4.85%
5. I find myself determining the final grade based on the overall impression a student leaves on me, without previous analysis of particular aspects of the oral presentation.	I completely disagree.	31	30.10%
	I mostly disagree.	37	35.92
	I am neutral.	16	15.53
	I mostly agree.	16	15.53%
	I completely agree.	3	2.91%
6. I have successfully evaluated oral presentations using standardized rating scales intended for assessing oral production activities.	I completely disagree.	36	34.95%
	I mostly disagree.	19	18.45%
	I am neutral.	14	13.59%
	I mostly agree.	23	22.33%
	I completely agree.	11	10.68%
7. I have attempted to evaluate oral presentations using standardized scales, but I find it challenging to understand and apply the descriptors for different proficiency levels.	I completely disagree.	51	49.51%
	I mostly disagree.	27	26.21%
	I am neutral.	11	10.68%
	I mostly agree.	12	11.65%
	I completely agree.	2	1.94%
8. I tend to evaluate different students' performances almost equally because it is challenging to determine why some students' presentations are more or less successful when compared to others.	I completely disagree.	43	41.75%
	I mostly disagree.	44	42.72%
	I am neutral.	7	6.80%
	I mostly agree.	8	7.77%
	I completely agree.	1	0.97%
9. I sometimes find it challenging to provide clear and precise feedback about students' oral presentation achievement.	I completely disagree.	57	55.34%
	I mostly disagree.	30	29.13%
	I am neutral.	11	10.68%
	I mostly agree.	2	1.94%
	I completely agree.	3	2.91%

The second statement was designed to check whether the oral presentation assessment was interesting and challenging. The teachers completely or mostly agree that it is an entertaining process (70.87%), while 14.56% of teachers completely or mostly agree that they do not find any satisfaction in assessing oral presentations, and the same number of teachers were neutral on this issue.

The results for the third statement confirm the previous findings, as the majority of teachers (63.17%) completely or mostly disagree that oral presentation assessment causes fatigue, while 22.33% of teachers mostly or completely agree that it is tiresome and exhausting work. As many as 16.50% of teachers could not make up their minds on this matter.

The fourth statement investigated whether the teachers had any difficulties deciding whether to evaluate different presentation features equally when making inferences about the final grade. Approximately one-half of respondents (53.40%) completely or mostly disagree with the statement, whilst others (46.60%) completely or mostly agree that they are undecided about how to evaluate different features of oral presentations. It thus turns out that a significant number of teachers admit to having some troubles when they need to evaluate different aspects of oral presentations.

The fifth statement was used to check whether the teachers reached any conclusions about the final grade based on the overall impression that a student left, without previous analysis of the individual aspects of their oral presentations. The results show that the greatest number of teachers (66.02%) mostly or completely disagree with this statement. The percentage of those who mostly or completely admit to evaluating their learners' performances based on the general impression they make is much lower, at only 15.53 %. Approximately the same number of teachers were undecided on this matter.

The analysis of the responses to the sixth statement shows that approximately half of the respondents (53.40%) do not evaluate their students' oral presentations based on the existing standardized rating scales. In contrast, 33.01% of the teachers do use the existing standardized scales. A minority of teachers (13.59%) were undecided on this question.

The seventh statement was designed to investigate whether the teachers tried to evaluate oral presentations based on the existing standardized scales while simultaneously having some difficulties understanding and using the descriptors for different levels of presentation proficiency. Most teachers stated that they completely or mostly agreed with the statement, i.e., they did not have trouble understanding and applying the descriptors to evaluate oral presentations. Only 13.59% of teachers stated that they have difficulties when assessing oral presentations according to the descriptors of the standardized scales, while 10.65% of the respondents were undecided.

The eighth statement checked if teachers occasionally evaluated different performances almost equally because it was difficult to decide why some presentations were successful when compared to others. The evidence shows that the greatest majority (84.47%) completely or mostly disagree with the statement, which means they feel that it is not difficult to decide how to evaluate different presentation characteristics appropriately. However, a small number

of teachers (8.74%) stated that they tend to give approximately the same grades to students because it is difficult to decide why some students' performances are more successful than others, and 6.80% of the teachers admitted that they were undecided on this issue.

The final statement aimed to reveal whether the teachers sometimes find it challenging to provide clear and precise feedback about students' oral presentations. For most of the teachers this procedure does not seem to be complicated, since 84.47% of them stated that they completely or mostly disagree with the ninth statement. A significantly smaller number of teachers (4.85%) find it challenging to give clear and precise feedback to their students, while 10.68 % were neutral on this issue.

The implication of the results presented above is that the teachers do not have difficulties when assessing LSP oral presentations. Furthermore, in general such assessments are not tiresome or exhausting, and can even be an interesting and challenging process. Additionally, the teachers do not have difficulties providing clear feedback to their students about their achievements in ways that can positively influence their future learning. However, they find it complicated to discern how to evaluate different aspects of an oral presentation, as well as how to evaluate the students' performances using already existing standardized scales, although they admitted to having understood and interpreted the standardized criteria and descriptors successfully when it was necessary.

The fourth part of the survey contained one open-ended question, the responses to which were analysed using content analysis. The teachers were asked to suggest recommendations for further improvement of the assessment process. A total of 84 teachers provided responses, which were then divided into 14 different thematic categories, as summarized in Table 3.

Content analysis of the responses indicates that most teachers believe that developing a specialized rating scale for assessing LSP oral presentations would be the most helpful tool for improving the assessment procedure. A similar number of respondents expressed interest in attending seminars, teacher training or workshops which would equip them with the knowledge and skills necessary for oral presentation assessment. Inter-rater discussions and the exchange of experience also featured as a helpful strategy to improve assessment. The teachers' suggestions also included some other useful recommendations: 1) introducing students to the rating process and providing detailed information about rating criteria, 2) allowing more time for assessment and detailed analysis of students' performances, 3) using standardized scales during assessment, 4) providing thorough preparation instructions to students before their final oral performances, 5) using formative instead of summative assessment, 6) equipping undergraduate student-teachers with the theoretical knowledge and skills necessary for the oral presentation assessment, 7) creating opportunities for teachers to explore relevant research studies, 8) producing specialized software which could facilitate the rating process, and 9) modifying the presentation topics to align with the content of specific subjects in the study programme. As we can see from the previous discussion, the need to generate a unique rating scale with clearly defined criteria and descriptors for LSP oral presentation assessment is highlighted as the most valuable strategy for assessing oral presentations.

TABLE 3. Recommendations for improvement of oral presentation assessment.

<b>Recommendations for improvement of oral presentation assessment</b>	<b>Number of responses provided by teachers</b>	<b>Percentage</b>
1. Creating a rating scale with clearly defined criteria for LSP oral presentation assessment.	19	22.62%
2. Organizing seminars, teacher training and workshops related to LSP oral presentation assessment.	17	20.24%
3. Inter-rater collaboration including discussions and the exchange of experience.	9	10.71%
4. Collaboration with experts from the field relevant to the LSP topic.	7	8.33%
5. Introducing students to the assessment process and informing them about assessment criteria.	6	7.14%
6. Enabling teachers to have more time for assessment and detailed analysis of students' performances.	5	5.95%
7. Using a standardized scale to assess oral presentations.	4	4.76%
8. Allowing teachers more time to instruct students on preparing for oral presentations.	4	4.76%
9. Recording oral presentations and subsequent analysis.	4	4.76%
10. Introducing formative assessment instead of summative assessment.	3	3.57%
11. Developing curricula for the undergraduate-level teaching methodology courses that students attend.	2	2.38%
12. Teachers' exposure to relevant studies and available literature on oral presentations.	2	2.38%
13. Creating specific software intended for oral presentation assessment.	1	1.19%
14. Connecting presentation topics with the content covered in professional course subjects.	1	1.19%

The previous findings can also provide the answers to our research questions. Namely, the greatest challenge for teachers when assessing an oral presentation is the evaluation of different rating criteria that describe specific aspects of an individual's presentation. In other words, determining the level of significance for each specific rating criterion is the most challenging aspect of assessing oral presentations. The teachers' answers to the second research question are also in line with this, since they state that the most effective way to improve the assessment process would be to generate a specific rating scale for LSP oral presentation assessment. We can also notice that rater training and inter-rater collaboration were also among the proposals that could assist in making the assessment process more efficient.

## 5 Conclusion

Giving an oral presentation is a common task in higher education (Girard, Pinar, and Trapp 2011; Alwi and Sidhu 2013). Being quite a complex and demanding task, it can prepare students for the future business environment, and can also help them build self-confidence and public speaking skills, which are required even after formal education has been completed (Alwi and Sidhu 2013, 98–99). Moreover, practicing public speaking prepares students to express creativity and develop critical thinking (Morita 2000, 289). Therefore, oral presentation represents a fruitful research topic, both as a pedagogic and assessment task. Scientific inquiry into oral presentation use at different levels of education may encompass a survey of a range of activities. This includes studies about optimal teaching approaches, instructions on crafting successful presentations, developing appropriate teaching materials, oral performance assessment, and beyond.

The main objective of the present paper was to reveal the most common practices of LSP teachers when assessing oral presentations. Another aim was to detect the greatest challenges and demands of the evaluation procedure, as well as to investigate what teachers consider to be the most helpful method to improve the assessment process and make it more efficient. To that end, we conducted research with 103 teachers who teach LSP in higher education institutions in Serbia and Croatia. The results of the survey analysis show that the teachers predominantly evaluate oral performances on the spot and independently, without another evaluator's assistance or an expert from the specific LSP field. They rarely attend rater training or professional education and development sessions connected with assessment. The assessment of oral presentations combines both summative and formative approaches. Most teachers assign numerical grades and give detailed feedback about the quality of students' performances, aiming to enhance further learning.

Regarding the needs and challenges of the assessment process, it appears that the teachers gladly take part in oral presentation assessment as an enjoyable activity that does not cause exhaustion, even if it lasts for a long time. However, the difficulty of establishing clear rating criteria and determining the level of significance for each criterion seems to be quite a challenge. Therefore, the teachers believe that designing a multicriteria rating scale with clear and comprehensive descriptors would significantly promote the assessment procedure. Rater training and development sessions, as well as collaboration with an expert from the specific LSP field, would be highly beneficial.

Among the limitations of the study are the choice of the research instrument, and the possibility that the respondents may have been disinterested or dishonest in their replies. It would thus be useful to apply different research instruments to check the results.

The findings of the study open some questions for further research, such as investigating the results of teacher-expert collaboration or designing teacher training sessions and workshops. Furthermore, the most valuable action would be developing a distinctive rating scale to make the assessment of oral presentations more objective and reliable.

## References

- Alshammari, Noor Aayed, and Abdelrahim Hamid Mugaddam. 2023. "Oral presentations as assessment criteria for EFL learners at university level: A Saudi instructors' perspective." *Eurasian Journal of Applied Linguistics* 9 (2): 132–41. <https://doi.org/10.32601/ejal.90201>.
- Alwi, Nor Fathiah Binti, and Gurnam Kaur Sidhu. 2013. "Oral presentation: Self-perceived competence and actual performance among UiTM Business Faculty students." *Procedia – Social and Behavioral Sciences* 90: 98–106.
- Atay, Derin, and Enisa Mede. 2017. "English language teachers' assessment literacy: The Turkish context." *Dil Dergisi* (1): 43–60. <https://www.cceol.com/search/article-detail?id=548906>.
- Basturkmen, Helen, and Catherine Elder. 2004. "The practice of LSP." In *The Handbook of Applied Linguistics*, edited by Alan Davies and Catherine Elder, 672–94. Malden: Blackwell Publishing.
- Bøhn, Henrik. 2015. "Assessing spoken EFL without a common rating scale: Norwegian EFL teachers' conceptions of construct." *SAGE Open* 5 (4): 1–12. <https://doi.org/10.1177/2158244015621956>.
- Brahim, Bellit, and Aliochouche Fadhila. 2021. "Speaking instruction in the Algerian secondary schools: EFL teachers' attitudes and classroom practices." *El Mohtaref Journal of Sports Science: Social and Human Science* 8 (1): 339–54. <https://www.asjp.cerist.dz/en/downArticle/380/8/1/154864>.
- Burkšaitienė, Nijolė. 2020. "Self-assessment of ESP oral performance: A tool for learning and fostering learner awareness." *The Journal of Education, Culture, and Society* 11 (1): 146–62.
- Council of Europe. 2001. *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*. Cambridge: Cambridge University Press. <https://rm.coe.int/16802fc1bf>.
- Fulcher, Glenn. 2012. "Assessment literacy for the language classroom." *Language Assessment Quarterly* 9 (2): 113–32. <https://doi.org/10.1080/15434303.2011.642041>.
- Girard, Tulay, Musa Pinar, and Paul Trapp. 2011. "An exploratory study of class presentations and peer evaluations: Do students perceive the benefits?" *Academy of Educational Leadership Journal* 15 (1): 77–94. <https://www.abacademies.org/articles/aeljvol15no12011.pdf>.
- Hill, Monica, and Anne Storey. 2003. "Speak Easy: Online support for oral presentation skills." *ELT Journal* 57 (4): 370–76. <https://doi.org/10.1093/elt/57.4.370>.
- Luoma, Sari. 2004. *Assessing Speaking*. Cambridge: Cambridge University Press.
- Morita, Naoko. 2000. "Discourse socialization through oral classroom activities in a TESL graduate program." *TESOL Quarterly* 34 (2): 279–310. <http://www.jstor.org/stable/3587953>.
- Norris, John M., James Dean Brown, Thom Hudson, and Jim Yoshioka. 1998. *Designing Second Language Performance Assessments*. Mānoa: University of Hawai'i.
- Palmour, Louise. "Assessing speaking through multimodal oral presentations: The case of construct underrepresentation in EAP contexts." *Language Testing* 41 (1): 9–34. <https://doi.org/10.1177/02655322231183077>.
- Petrović, Vesna, Ana Radović Firat, and Lidija Palurović. 2022. "Oral presentation as authentic material for providing ESP instruction." In *9th International Scientific Conference Technics and Informatics in Education – TIE 2022*, edited by Ivan Miličević, 414–19. Čačak: Faculty of Technical Sciences. <https://doi.org/10.46793/TIE22.414P>.
- Polovina, Nina, and Tanja Dinić. 2014. "Načini ocenjivanja studentskih prezentacija u nastavi stranog jezika struke." In *Zbornik radova. Treća međunarodna konferencija Strani jezik struke: prošlost, sadašnjost, budućnost*, edited by V. Caklejić, A. Vujović and M. Stevanović, 409–16. Beograd: Fakultet organizacionih nauka.
- Richards, Jack C. 2008. *Teaching Listening and Speaking: From Theory to Practice*. Cambridge: Cambridge University Press.
- Rizvi, M. Ashraf. 2004. "Student self-assessment in ESP: A study of students' learning perceptions." *HOW* 11 (1): 75–86. <https://www.howjournalcolombia.org/index.php/how/article/view/180/229>.
- Smit, Eva. 2020. "Assessing oral proficiency in the EFL classroom: A qualitative study of teachers' understanding, experience, and assessment of oral production, and interaction in Swedish upper secondary schools." Bachelor's Thesis, Linnaeus University. <https://www.diva-portal.org/smash/get/diva2:1604147/FULLTEXT01.pdf>.

- Vangelisti, Anita L., John A. Daly, and Gustav W. Friedrich. 1999. *Teaching Communication*. Mahwah: Lawrence Earlbaum Publishers.
- Vogt, Karin, and Dina Tsagari. 2014. "Assessment literacy of foreign language teachers: Findings of a European study." *Language Assessment Quarterly* 11 (4): 374–402. <https://doi.org/10.1080/15434303.2014.960046>.
- Yildiz, Lilli Mari. 2011. "English VG1 level oral examinations: How are they designed, conducted and assessed?" Master thesis, University of Oslo.





# **Part IV**

## **Translation and Contrastive Studies**



# A Corpus-Based Contrastive Analysis of Metaphorical Collocations of the Terms *Right* and *Pravo*: Do Different Legal Cultures Give Rise to Different Metaphors?

## ABSTRACT

The paper focuses on the contrastive analysis of metaphorical collocations of the nouns *right* and *pravo* in the web corpora of English and Croatian. It categorizes metaphorical collocations based on the Master Metaphor List (Lakoff, Espenson, and Schwartz 1991), and attempts to detect the similarities and differences in the cognitive structuring of the two terms. The paper is thus an attempt to investigate the metaphorical dimension of legal collocations, a perspective largely ignored to date in the study of collocations in legal discourse. By combining the corpus-based approach, which allows us to access a significant number of collocations, and using the Master Metaphor List, the paper reveals the network of conceptual mappings operating in the background of collocations and proposes that such cognitive structuring cuts across different legal cultures.

**Keywords:** Croatian, English, legal phraseology, metaphorical collocations, metaphorical images, web corpora

## Korpusna kontrastivna analiza metaforičnih kolokacij izrazov *right* in *pravo*: ali različne pravne kulture ustvarjajo različne metafore?

## IZVLEČEK

Prispevek se osredotoča na kontrastivno analizo metaforičnih kolokacij samostalnikov *right* in *pravo* v spletnih korpusih angleščine in hrvaščine. Metaforične kolokacije kategorizira na podlagi seznama Master Metaphor List (Lakoff et al. 1991) ter poskuša odkriti podobnosti in razlike v kognitivni strukturiranosti obeh izrazov. Članek je tako poskus raziskovanja metaforične razsežnosti pravnih kolokacij, ki je bila doslej pri preučevanju kolokacij v pravnem diskurzu večinoma prezrta. S kombinacijo korpusnega pristopa, ki nam omogoča dostop do precejšnjega števila kolokacij, in z uporabo glavnega seznama metafor članek razkriva mrežo konceptualnih preslikav, ki delujejo v ozadju kolokacij, in predlaga, da takšna kognitivna strukturiranost preči različne pravne kulture.

**Ključne besede:** hrvaščina, angleščina, pravna frazeologija, metaforične kolokacije, metaforične podobe, spletni korpusi

# 1 Introduction<sup>1</sup>

While it is well-known that there have been only very few attempts in legal phraseology to propose a typology of phraseological units (Kjær 1990; 2007; Biel 2014; Goźdz-Roszkowski 2011; Kopaczyk 2013), most scholars suggest that the language of law does not abound in idiomatic expressions. Instead, its phraseology mainly includes multi-word terms, specialized collocations and formulaic expressions (Kjær 1990), but also lexical bundles (Goźdz-Roszkowski 2011; Kopaczyk 2013), i.e., recurrent word combinations fulfilling specific functions in legal discourse that are best revealed through a corpus-driven approach (Goźdz-Roszkowski and Pontrandolfo 2015, 133). The most recent typology of phraseological units in legal discourse is based on Kjær's (1990, 2007) and Goźdz-Roszkowski's (2011) research, and includes text-organizing patterns (e.g., amending and closing formulas); grammatical patterns (e.g., expressing obligation with the modal verb *shall*); term-forming patterns (i.e., multi-word terms); term-embedding collocations (e.g., *to hold shares*); and lexical collocations (e.g., *subject to this Regulation*) (Biel 2014).

This paper focuses on term-embedding collocations of the nouns *right* in English and *pravo* in Croatian, and suggests that, although such combinations are not idiomatic in nature, they do owe their existence to conceptual metaphors. This also seems natural given the central role that a metaphor plays in legal reasoning (Lakoff and Johnson 1980, 268; Winter 2001). Since the metaphorical dimension of collocations has been either ignored or not well-defined thus far (Patekar 2022; Stojić and Košuta 2022), both in phraseology of general-purpose (LGP) and special-purpose language (LSP), the project *Metaphorical Collocations – Syntagmatic Word Combinations between Semantics and Pragmatics* was launched at the University of Rijeka to investigate this subtype of collocations in the web corpora of the English, German, Croatian and Italian languages. Some authors refer to such combinations as metaphorical collocations (Reder 2006; Volungevičienė 2008; Konecny 2010, 2012), and suggest that they are characterized by a certain degree of stability rooted in the semantic cohesion between the constituents of such combinations, which in turn results from the process of metaphorization or metonymization (Stojić 2024, 11). Although it seems that metaphors and metonymies contribute to the stability of collocations, the fact remains that the higher the degree of stability between the constituents, the more faded the metaphor (ibid.). Consequently, we sometimes need to resort to the diachronic approach to the analysis of such combinations in order to detect the metaphorical concept operating in their background. Furthermore, such an approach will have to include the analysis of subject-specific knowledge in order to be able to explain the specific meaning that a collocate produces when coupled with a legal term.

In legal discourse, the presence of metaphors was recognized as early as in 1816, when Jacob Grimm published his paper *Von der Poesie im Recht* (Grimm 1816), indicating that the multinomials prevailing in both law and poetry have the same origin. His claim, however, was challenged by Kopaczyk (2017), whose research suggests that the usage of multinomials has significantly increased over time. With the exception of multinomials, the metaphorical

---

<sup>1</sup> This work has been fully supported by the Croatian Science Foundation within the project “Metaphorical collocations – syntagmatic word combinations between semantics and pragmatics” (IP-2020-02-6319).

dimension of legal phrasemes has been largely ignored hitherto in the research on legal phraseology. As a matter of fact, although metaphors have received significant attention in other specialized discourse, such as economics (Herrera-Soler and White 2012) and politics (Goatly 2007), “the interface of cognitive metaphor studies and legal studies is still a relatively novel and largely uncharted territory” (Wojtczak and Witczak-Plisiecka 2019, 275). Other than Winter’s research on the role of metaphor in legal reasoning (2001), two studies that made an overview of metaphors in different types of legal texts (Imamović 2013; Richard 2014), two pieces of research on the use of metaphorical expressions in Polish legislation (Wojtczak and Witczak-Plisiecka 2019) and legal English textbooks (Kordić 2023), there are no systematic studies on metaphorical collocations in legal language, let alone ones that adopt a corpus-based perspective. The paper thus attempts to fill this gap by suggesting that the terms *right* and *pravo* and the contexts in which they are embedded tend to be largely metaphorical.

In order to prove this claim the paper draws on the insights of Conceptual Metaphor Theory (CMT) and tries to group metaphorical collocations according to the Master Metaphor List (Lakoff, Espenson, and Schwartz 1991),<sup>2</sup> an extensive inventory of conceptual metaphors that includes four main general metaphorical systems: event structure, mental events, emotions and others. The inventory represents the first, hierarchically ordered list of metaphors, with the four general metaphorical systems including further specifications and sub-specifications. According to Lakoff and Johnson, a conceptual metaphor enables us to understand one thing in terms of another (1980, 5). It is a cognitive ability that manifests itself on a linguistic level (Ostroški Anić 2019, 41), and metaphorical collocations as phraseological units thus seem to owe their existence to the fact that the whole human conceptual system is largely metaphorical (Lakoff and Johnson 1980, 6). Furthermore, metaphorical collocations represent metaphorical concepts that are arranged systematically, and, as suggested by Lakoff and Johnson, by studying such expressions we can get an insight into the “coherent system” (1980, 9) of metaphorical concepts and the metaphorical expressions that they are represented by in the language itself. Such systematicity, as will be shown in this paper, will result in multiple metaphorical images for one specific collocation, confirming the claim that concepts can be analysed from different perspectives relative to our experience (Lakoff and Johnson 1980, 76). A careful study will, however, reveal that one concept is more basic than the other, or is superordinate to the other, making the introduction of other concepts superfluous. This in turn allows us to arrive at a network of metaphorical images and group more specific images into one superordinate, which might provide aid in assigning concepts to metaphorical collocations of other nouns sharing the same or a similar semantic field.

## 2 Metaphorical Collocations

Since it is the task of every phraseologist to specify the object of their research, “in order to (i) render their definitions maximally precise and (ii) allow researchers from other frameworks

<sup>2</sup> Although fully aware of the fact that the list produced by Lakoff, Espenson and Schwartz (1991) is far from complete, and that in the last three decades new repositories have been designed (e.g., the MetaNet project at ISCL, Berkeley), the project team opted for the Master Metaphor List due to the fact that the MetaNet repository was at that time migrating and was not available from the main webpage (<https://metanet.arts.ubc.ca/metaphor-databases>).

to more easily recognize potential areas of overlap, or indeed conflict” (Gries 2008, 3), we will at this point take up this challenge and embark upon defining metaphorical collocations. Approaching this task, however, entails laborious study, for there is to date no consensus on a widely accepted definition of collocations (Patekar 2022), let alone metaphorical ones. As a matter of fact, Gries’ proposal of six criteria of phraseologisms (2008) suggests that the nature and number of the elements in a phraseologism, the number of times an expression must occur to be regarded as a phraseologism, the permissible distance between the elements of a phraseologism, the degree of lexical and syntactic flexibility of elements, the role of the semantic unity, and semantic non-compositionality are largely dependent on researcher’s input. Furthermore, the advance of corpus linguistics has widened the concept of the collocation, which now also includes more extended units of meaning (Sinclair 2004), or “extended collocations”, i.e., “two-word combinations augmented by at least one lexical item” (Gabrovšek 2014, 11). Some studies even suggest that the base of the collocation can in some contexts, especially in specialized discourse, be extended to include more than one element (Dobrić Basanež 2017; 2018). It thus goes without saying that anyone who tries to precisely define collocations is at risk of receiving criticism for either being too narrow or too wide in their view. That said, and fully aware of the potential criticism, we will approach our definition of collocations both from the frequency-based and phraseological approaches (Nesselhauf 2004), and construe the notion of collocations to include only binary collocations, with their minimum relative frequency set at five per million words. The metaphorical dimension, on the other hand, is another complicating factor, mostly due to the fact that there is again no consensus among researchers on what is or should be metaphorical in a collocation. Some previous research has shown that the metaphorical dimension is usually reflected by the collocate (Stojić and Murica 2010; Stojić and Štiglić 2011; Stojić and Barić 2013; Stojić 2015). Very often the metaphorical potential of the collocate is too weak to contribute to the figurativeness of the collocation as a whole, and remains only vaguely distinguishable in the collocate component, thus resulting in a “faded metaphor” (Patekar 2022). Our understanding of metaphorical collocations is closest to that of Philip (2011), who compares metaphorical collocations to idioms, but suggests that the main distinction between the two lies in the fact that in metaphorical collocations the meaning of their individual components remains transparent. Moreover, the results of our project have shown that in polysemous nouns the process of metaphorization occurs both at the level of the base and that of the collocate (Stojić and Matešić 2024). This also holds true for the nouns analysed in this paper, since both *right* and *pravo* by their etymology refer to something straight<sup>3</sup> or stretched out.<sup>4</sup> It must be pointed out, however, that in specialized discourses the concept of metaphorical collocations includes both terminological and non-terminological collocations, but this paper focuses only on terminological ones and disregards the concordances in which the collocates display their non-specialized meaning (e.g., *to be right*).

### 3 Corpus and Methodology

As addressed elsewhere in this paper, the corpora used for extracting metaphorical collocations are enTenTen20 for English and hrwac 2.2. for Croatian, the only available and

<sup>3</sup> See [https://hjp.znanje.hr/index.php?show=search\\_by\\_id&id=eVhnXhc%3D](https://hjp.znanje.hr/index.php?show=search_by_id&id=eVhnXhc%3D). Accessed June 10, 2023.

<sup>4</sup> See <https://www.etymonline.com/search?q=right>. Accessed June 10, 2023.

representative web corpus for the Croatian language. Since the paper deals with collocations of the terms *right* and *pravo*, we extracted collocational profiles of the selected nouns by means of the wordsketch function. The following grammatical relations remain relevant for our study: “modifiers of x”, “adjective predicates of x”, “verbs with x as object”, “verbs with x as subject”, and “possessors of x”. The equivalent relations in Croatian include “kakav?”, “biti kakav?”, “particip”, “subjekt\_od”, “objekt\_od” and “n-koga čega”. It is obvious that for Croatian there is an additional grammatical relation, given the fact that in the English corpus the participle structure is realized throughout four categories.<sup>5</sup> Furthermore, enTenTen20 suggests that the relation “possessors of x” is also rather repetitive or irrelevant for English nouns, given that the “modifier of x” structure also includes noun+noun combination. For this reason, there are only four salient grammatical relations worth studying in the English corpus. A further step in the methodology included careful study of the possible collocation candidates in Excel, whereupon we indicated whether a combination displayed through wordsketches is a collocation, and if, yes, whether it is metaphorical or not. Metaphorical collocations were identified by means of the MIP method (Pragglejaz Group 2007), which allows researchers to distinguish between the basic and contextual meaning of lexical units. The method proposes that collocations should be analysed constituent by constituent, which is why we first analysed the meaning of the base and then turned to the analysis of the collocates. Since both *right* and *pravo* are terms, their meaning was analysed by turning to legal dictionaries. As for the collocates, we used the *Merriam Webster Dictionary*<sup>6</sup> and *Online Etymological Dictionary*<sup>7</sup> for the English collocations and *Hrvatski jezični portal* (HJP)<sup>8</sup> for the Croatian candidates. The final step involved a classification of metaphorical collocations according to the networks of conceptual metaphors proposed in the Master Metaphor List (Lakoff, Espenson, and Schwartz 1991), and a comparison of the results from both corpora to reveal the similarities and differences between the mappings in two different legal cultures. The metaphors outlined in this paper are mostly taken directly from this inventory, although some are further elaborated based on the analogy with some other abstract nouns. In Lakoff, Espenson and Schwartz (1991) rights are coupled with possessions, although not monetary ones, but money is, on the other hand, portrayed as a liquid, which justifies our reasoning for the image RIGHT IS A LIQUID MONETARY POSSESSION. Similarly, in the OTHER group, we assumed that, if obligations are viewed as children in Lakoff, Espenson and Schwartz (1991) and are subsumed under OTHER, then we must also perceive rights in the same way, as evidenced by the comparison between the following examples:

*He's been neglecting his duties.* (Lakoff, Espenson, and Schwartz 1991, 205).

*Sva njegova politička prava su zanemarena.* (hrwac, zadarskilist.hr)

<sup>5</sup> Sketch Engine very often lists –ing and –ed participles under the relation “modifiers of x” (e.g., *bargaining rights*) or “adjective predicates of x” (e.g., *rights are respected*), or even falsely recognizes a particular relation (e.g., in the combination *eating rights* the noun *right* is obviously not the object of *eat*, instead, *eating* serves as a modifier of *rights*).

<sup>6</sup> See <https://www.merriam-webster.com/>. Accessed 10 April 2024.

<sup>7</sup> See <https://www.etymonline.com/>. Accessed 10 April 2024.

<sup>8</sup> See <https://hjp.znanje.hr/>. Accessed 10 April 2024.

## 4 Results

The results shown in this section suggest that law is largely metaphorical, despite the fact that certainty of expression is often cited as the objective of legal language. Recent research has shown that in law there seem to be four major metaphor groups coupling legal concepts with natural elements, war, and different objects or persons (Richard 2014, 4). Some of these categorizations can be challenged, since the fact that a collocate invokes a certain image – for instance, that of an object – does not imply the existence of the same image in the collocation as the lexical unit. A case in point is the collocation *to break the law*, which might lead us to believe that law is perceived as an object that can be broken. As argued by Winter (2001, 14), we do not destroy the law here, we simply transgress its limits. If this were not so, the language would abound in expressions indicating the repair of the law. Their non-existence, however, gives rise to the conceptual metaphor *LAW IS A CONSTRAINT*, and is further supported by collocations like *law-abiding* or *get around the law* (Winter 2001, 14). Winter's findings thus propose that in order to avoid misconceptions in the classification of image schemas, we have to consider a significant number of semantically related collocates centring around one term, which in turn advocates a corpus-based approach. Furthermore, the grouping of the most frequent metaphorical collocations according to the Master Metaphor List (Lakoff, Espenson, and Schwartz 1991) assists in producing a network of interrelated general and specified images, which without such a combined approach might go unnoticed. Although the analysis, as can be seen from the examples in the next section, is mostly focused on the semantic meaning of collocations, concordances also reveal the morphological behaviour of the nouns *right* and *pravo*, suggesting that alternation is possible between singular and plural forms. Additionally, when the noun *right* occurs in singular form, the collocate is usually preceded by the definite article (e.g., *the basic / fundamental / universal right*).

### 4.1 Metaphorical Collocations of the Noun *Right*

The noun *right* has multiple meanings, but its prototypical legal sense includes “an entitlement to something, whether to concepts like justice and due process or to ownership of property or some interest in property, real or personal”<sup>9</sup>. The process of metaphorization underlying collocations in this section is thus the classical one of coupling abstract concepts with the concrete, and most images belong to the event structure group.

#### EVENT STRUCTURE

PROPERTIES ARE PHYSICAL PROPERTIES

IMPORTANCE IS A PHYSICAL PROPERTY

IMPORTANCE IS INTERIORITY / CENTRALITY

Modifiers of x: basic / fundamental right

Adjective predicates of x: right is basic / applicable / integral

Verbs with x as object: constitute / uphold / support / undermine rights

---

<sup>9</sup> <https://dictionary.law.com/Default.aspx?typed=right&type=1>. Accessed 5 June 2023.

## AMOUNT IS VERTICALITY

Modifiers of x: collective rights

Verbs with x as subject: rights accrue

## RIGHTS ARE POSSESSIONS

Modifiers of x: inalienable / inherent / indispensable / vested right

Verbs with x as object: own / hold / retain / possess / obtain / maintain / enjoy / relinquish the right

## RIGHTS ARE MONETARY POSSESSIONS

Verbs with x as object: acquire / grant / earn / gain / reserve / sell / buy / purchase / confer / assign / transfer rights

## RIGHT IS A CONTAINER

Modifiers of x: full rights

Adjective predicates of x: right is intact / void / limited / broad

Verbs with x as object: expand / extend rights

## GAINING AND MAINTAINING RIGHTS IS FIGHTING A WAR

Adjective predicates of x: right is enforceable

Verbs with x as object: assert / claim / enforce / defend / surrender / safeguard / protect / violate / forfeit / infringe / win / lose / champion / respect / restore the rights

Verbs with x as subject: rights trump

being better in a static situation is being above (more is up / high status is up)

Modifiers of x: divine / absolute / universal right

Adjective predicates of x: right is superior / sacred / sacrosanct

Verbs with x as subject: rights allow / permit

## CHANGE IS MOTION / PROGRESS IS FORWARD MOTION

Verbs with x as object: terminate the right

Verbs with x as subject: rights expire

## ACTION IS SELF-PROPELLED MOTION

Adjective predicates of x: right is exercisable

Verbs with x as object: exercise rights

## OTHER

RIGHT IS A PAINTER

Verbs with x as subject: rights illustrate / depict

SOCIETY IS A BODY

Adjective predicates of x: right is vital / weak

Verbs with x as subject: rights rest

## 4.2 Metaphorical Collocations of the Noun *pravo*

The Croatian noun *pravo* is a homonym and its different meanings are sometimes separated in dictionaries, indicating its different accents. Thus, *pràvo* refers to something which is right, proper, or equitable, whereas the other form *právo* is polysemous and includes two related meanings, those of 'law' and 'right'.<sup>10</sup> Most meanings displayed by the collocates in Sketch Engine, however, are equivalents of the English noun *right*, with the exception of the image MORE / HIGH STATUS IS UP. The largest portion of metaphorical collocations therefore correspond to the images found in enTenTen20.

### EVENT STRUCTURE

PROPERTIES ARE PHYSICAL PROPERTIES

IMPORTANCE IS A PHYSICAL PROPERTY

IMPORTANCE IS INTERIORITY / CENTRALITY:

Kakav?: osnovano / elementarno pravo

Biti kakav?: pravo je utemeljeno / zasnovano / temeljeno / ustanovljeno

PROPERTIES ARE POSSESSIONS

RIGHTS ARE POSSESSIONS

Kakav?: zajamčeno / zagarantirano / oduzeto pravo

Biti kakav?: pravo je pridržano / zadržano

n-koga-što?: oduzeti / uzimati / uskratiti / uživati pravo

n-koga-čega?: nositelj prava

RIGHTS ARE (LIQUID) MONETARY POSSESSIONS

Biti kakav?: pravo je dodijeljeno / prodano

n-koga-što?: polagati / stjecati pravo

n-koga-čega?: izvor prava

---

<sup>10</sup> Hrvatsko-engleski rječnik prava, 2019, p. IX, Zagreb: Narodne Novine.

## RIGHT IS A CONTAINER

Kakav?: puno pravo

Biti kakav?: pravo je ograničeno / isključeno / prošireno / suženo

n-koga-čega: okvir / opseg prava

n-koga-što?: suziti pravo

## GAINING AND MAINTAINING RIGHTS IS FIGHTING A WAR

Kakav?: procesno / zaštićeno pravo

Biti kakav?: pravo je izboreno / prekršeno / utvrđeno / narušeno

n-koga-što?: osporavati / ugrožavati / štititi pravo

n-koga-čega?: obrana / zlouporaba prava

## HARM IS DESTRUCTION

Particip: pogaženo / zakinuto pravo

n-koga-čega?: rezanje prava

## HARMING IS LOWERING / HELPING IS RAISING / EQUALIZING

Biti kakav?: pravo je umanjeno / uvećano / izjednačeno

n-koga-čega?: jednakost prava

## HARM IS PREVENTING FORWARD MOTION TOWARD A GOAL

## HARM IS AN OBSTACLE

Kakav?: mjerodavno / obvezno

Biti kakav?: pravo je propisano

## BEING BETTER IN A STATIC SITUATION IS BEING ABOVE (MORE IS UP / HIGH STATUS IS UP)

Kakav?: univerzalno / apsolutno pravo

## CHANGE IS MOTION / PROGRESS IS FORWARD MOTION R

Subjekt\_od: pravo prestaje

n-koga-čega: trajanje prava

**OTHER**

## RIGHT IS A CHILD

Kakav?: povrijeđeno pravo

Particip: pravo je ostavljeno / iskorišteno / zadovoljeno / zanemareno

## RIGHT IS AN EMPLOYEE

Particip: suspendirano pravo

## 5 Discussion

Upon comparing the different categories of metaphorical collocations presented in the previous sections, it seems important to point out that in both English and Croatian legal language rights are perceived as possessions, or abilities to possess. Although this image does occur in the Master Metaphor List (Lakoff, Espenson, and Schwartz 1991), it had not been fully developed yet when the inventory was published, as pointed out by the authors. Data from our corpora thus give significant proof for the existence of the image RIGHTS ARE POSSESSIONS and its different specifications. In the English corpus rights are perceived as something that cannot be taken away from a person (e.g., *inalienable / inherent / indispensable rights*), and that sometimes have to get covered or dressed in robes in order to achieve that (e.g., *vested right*), whereas in the Croatian corpus there is an invocation of a formal assurance (e.g., *zajamčeno / zagarantirano pravo*) as a guarantee that a right cannot be taken away (e.g., *uzimati / oduzeti pravo*). In both corpora rights also seem to be portrayed as monetary possessions (e.g., *acquire / grant / earn / gain / reserve / sell / buy / purchase / confer rights; pravo je dodijeljeno / prodano; polagati pravo*), but in the Croatian corpus rights are additionally portrayed as liquid monetary possessions (e.g., *stjecati pravo; izvor prava*). Possessing rights is in both corpora perceived as holding or carrying them (e.g., *hold / retain / possess / obtain the right; pravo je pridržano / zadržano / nositelj prava*), but the image POSSESSING IS HOLDING was not introduced as an additional specification, given the fact that the difference between some collocates belonging to this group is rather subtle (e.g., *own* and *possess*) and is no longer visible today (e.g., *possess* – in the late 14th century referred to “holding” without regard to ownership;<sup>11</sup> *obtain* comes from the Latin *obtinere*, “hold, hold fast”<sup>12</sup>). A further argument in favour of not introducing this image separately from the neutral RIGHTS ARE POSSESSIONS image is the fact that the verb *to hold* is very often used as an antonym of the verb *to buy* in the context of property. Furthermore, the binomial *to hold and enjoy the property* can be reduced to a single verb *enjoy*, suggesting that enjoying property includes possessing and holding it (Dobrić Basanež 2017, 159).

Another prevalent image in both languages is that of GAINING AND MAINTAINING RIGHTS IS FIGHTING A WAR, which accounts for numerous collocations (e.g., *defend / surrender / safeguard / protect / win / lose / champion the rights; pravo je izboreno / prekršeno / utvrđeno / narušeno; obrana prava*). Some units from this group might also induce the image of sports (e.g., *champion the rights*) or the verticality schema (e.g., *rights trump*), but since the collocate *champion* implies the action of a militant supporter of something,<sup>13</sup> and *trump* has been linked to *triumph*,<sup>14</sup> which in turn induces the meaning of winning as if by military force,<sup>15</sup> the decision that they belong to the war schema seemed reasonable. Other collocations from this group might invoke the metaphor RIGHT IS A BUILDING. A case in point may be illustrated with the collocation *utvrđeno pravo*, which gives the image of a fortress, but since we are well aware of the fact that such fortresses represent military strongholds, the image of a building

<sup>11</sup> See <https://www.etymonline.com/search?q=possess>. Accessed February 5, 2024.

<sup>12</sup> See <https://www.etymonline.com/search?q=obtain>. Accessed February 5, 2024.

<sup>13</sup> See <https://www.merriam-webster.com/dictionary/champion>, Accessed April 10, 2024.

<sup>14</sup> See [https://www.etymonline.com/word/trump#etymonline\\_v\\_17893](https://www.etymonline.com/word/trump#etymonline_v_17893). Accessed April 10, 2024.

<sup>15</sup> See <https://www.merriam-webster.com/dictionary/triumph>. Accessed April 10, 2024.

seems to be rather insignificant here. This can be supported by the collocate *utvrđeno* that stems from the adjective *tvrd*, whose meaning, among other things, implies being “resistant to the usage of weapon”.<sup>16</sup> Similarly, although the collocations *to violate a right* and *prekršeno pravo* might belong to the image HARM IS AN OBSTACLE, they differ from the collocation *to break the law*, since, the verb *violate* does include a certain degree of violence<sup>17</sup> in the same way that the participle *prekršen* refers to being broken with force.<sup>18</sup> In addition, in the Croatian corpus there are several images of harm, which seem to be non-existent in English. One might argue that the collocates representing the image HARM IS DESTRUCTION (e.g., *pogaženo / zakinuto pravo; rezanje prava*) could also be grouped into the mapping IDEAS ARE PLANTS, since they all imply the meaning of stepping on or cutting down of a plant. However, given the fact that there are no antonymous pairs (e.g., *rights blossom / sprout*), their allocation to the negatively connoted image of destruction seemed justified. Another negative image of harm is the one where harm is coupled with the lowering of rights and is complemented with the image HELPING IS RAISING / EQUALIZING, which in turn invokes a positive semantic prosody in collocates *jednakost* and *uvećati*. Finally, the image HARM IS AN OBSTACLE only includes collocates where the noun *right* refers to ‘laws’ or ‘rules’.

In both corpora rights are perceived as containers, either full (e.g., *full rights; puno pravo*) or void (e.g., *void rights*), with the defined framework or scope (e.g., *opseg / okvir prava*) that can be extended or narrowed (e.g., *proširiti / suziti prava; extend rights / limited rights*). Some of these collocations could also belong to the concept of building as well, especially if we consider that the content of the container image is situated in its interior, whereas in the building concept, the content is its foundation and the outer shell (Lakoff and Johnson 1980, 100). But since the container and the building metaphor are corresponding images, they can be freely mixed (Lakoff and Johnson 1980, 100). This also applies to the image where the physical property of importance is coupled with centrality or depth, given the fact that in the building metaphor the ground level is the surface that defines both the content and the depth (Lakoff and Johnson 1980, 100). The image IMPORTANCE IS CENTRALITY, however, seems more reasonable here, since when coupled with the noun *right* all these collocates invoke the meaning of being most important.

The image of more being up is also present in both corpora and is motivated by the verticality schema. Universal and absolute rights are thus rights that should belong to all, but this universality and absoluteness is culturally dependent, suggesting that some rights are more universal (e.g., *the right to life*) than the others (e.g., *the right to marry*) and are thus superior to them. Similarly, a divine right used to belong to people in power, who had control over ordinary citizens, as evidenced by the following concordance from the corpus:

*As early as the 7th century A.D., the Anglo-Saxon kings of England claimed to rule by divine right; to disobey them was to disobey God.* (hnn.us)

Rights themselves are also given the position of control or authority and are “up” when they allow or permit something. There is a slightly more static image of verticality visible in the collocates *extreme / collective rights* and *rights accrue*, where the amount is associated

<sup>16</sup> See [https://hjp.znanje.hr/index.php?show=search\\_by\\_id&id=f19iWRh4](https://hjp.znanje.hr/index.php?show=search_by_id&id=f19iWRh4). Accessed January 20, 2024.

<sup>17</sup> See <https://www.etymonline.com/search?q=violate>. Accessed January 20, 2024.

<sup>18</sup> See [https://hjp.znanje.hr/index.php?show=search\\_by\\_id&id=elhVWhY%3D](https://hjp.znanje.hr/index.php?show=search_by_id&id=elhVWhY%3D). Accessed January 20, 2024.

with the vertical increase in size. The increase is at times rather neutral (e.g., *rights accrue*) and sometimes it implies danger (e.g., collective rights emerging as a result of insufficient protection by individual human rights<sup>19</sup>). Both corpora also reveal the image of change and relative motion, as evidenced by collocates denoting the time dimension (e.g., *terminate rights*; *rights expire*; *trajanje prava / pravo prestaje*).

Finally, in both corpora rights are portrayed as persons, but since personification as a subtype of ontological metaphors (Lakoff and Johnson 1980) is a general category, this paper dissects different aspects of persons, and reveals that rights are in the English corpus perceived as painters (e.g., *rights illustrate / depict*), while in the Croatian corpus they are presented as children (e.g., *ostavljeno / povrijeđeno pravo*) or employees (e.g., *suspendirano pravo*). These images are not present in the Master Metaphor List, but since Lakoff, Espenson and Schwartz (1991) list OBLIGATIONS ARE CHILDREN as one of conceptual metaphors in the OTHER group, we opted for the same general schema. In the English corpus personifications also account for the SOCIETY IS A BODY metaphor, where rights are portrayed as vital, weak or resting on something.

The analysis in this paper has revealed that many collocates listed in the above sections are already lexicalized and we do not tend to consider the combinations into which they enter as metaphorical. The etymology, however, contradicts this claim and suggests, *inter alia*, that the component of “being sticky” is no longer present in the collocate *inherent*, but is instead replaced by the mental picture of being part of someone’s constitution or belonging to them by nature.<sup>20</sup> Similarly, we do not perceive the collocate *vest* as something that necessarily refers to clothes (from the Old French *vestir*, “to clothe; get dressed”<sup>21</sup>), but as something which entails the placing of rights into someone’s possession. In the same vein, today neither *neotuđiv* nor *inalienable* include the meaning of belonging to the other (which was the meaning of the Latin word *alius*, via which the word entered the English language), but are used strictly as legal terms, referring to the impossibility of being transferred to a third person. The collocation *narušeno pravo* no longer involves the mental picture of destroying a building in the same way, as *obrana prava* includes no physical defence or presence of the army. Nevertheless, although we might not perceive these collocates as metaphorical, there seems to be a cognitive structuring behind these combinations, which, as suggested above, very often cuts across cultures. The universality or near-universality (Kövecses 2006, 155) of metaphors between English and Croatian comes best to the foreground with event structure metaphors, where the only difference seems to be detected in the image of harm. It thus seems that, although the legal languages that were under scrutiny in this paper belong to different legal cultures, namely, common law and civil law, they view rights in the same way. Finally, it appears that the impact is more far-reaching than one tends to think, as it regulates the systematicity of metaphorical images, which in turn affects the semantic productivity of metaphorical collocations.

---

<sup>19</sup> See <https://www.foei.org/what-are-collective-rights/>. Accessed April 10, 2024.

<sup>20</sup> See <https://www.merriam-webster.com/dictionary/inherent>. Accessed June 18, 2023.

<sup>21</sup> See <https://www.lexisnexis.co.uk/legal/glossary/applicable-law>. Accessed June 11, 2023.

## 6 Conclusion

The contrastive analysis of metaphorical collocations of the nouns *right* and *pravo* conducted for the purpose of this paper suggests that the Croatian and English legal cultures mostly perceive these concepts in the same way, as witnessed by the similarities between the detected conceptual mappings. Furthermore, the mappings suggest that both nouns enter into combination with collocates that refer to events. What is more, although the differences in terminology and phraseology between the two legal languages may be vast, it seems that the conceptual images operating in their backgrounds do not tend to differ significantly.

The paper is an attempt to show that by combining a corpus-based approach and assigning conceptual metaphors to the extracted collocations, we can achieve a systematicity of metaphorical images. Furthermore, although the concepts operating in the background of metaphorical collocations can be analysed in more than one way, the inventory of semantically related collocates extracted from corpora is crucial in determining the basic mapping. Needless to say, the analysis in the paper is far from flawless and the concepts discussed in it are no “ultimate building blocks” (Lakoff and Johnson 1980, 76), and can be analysed further. Future research might therefore profit from the data contained in specialized legal corpora and include a contrastive analysis of the nouns *law* and *zakon* to provide a complete inventory comparable to all collocates of the polysemous noun *pravo*. Nonetheless, this study’s investigation of the metaphorical dimension of collocations based on comparable corpora marks an innovation in the study of legal phraseology, and can possibly represent a further strand of research within it.

## References

- Biel, Łucja. 2014. *Lost in the Eurofog: The Textual Fit of EU Translated Law*. Frankfurt am Main: Peter Lang.
- Dobrić Basanež, Katja. 2017. “Extended binomial expressions in the language of contracts.” In *Phraseology in Legal and Institutional Settings: A Corpus-Based Interdisciplinary Perspective*, edited by Stanisław Goźdz-Roszkowski and Gianluca Pontrandolfo, 203–20. London: Routledge.
- Gabrovšek, Dušan. 2014. “Extending binary collocations: (Lexicographical) implications of going beyond the prototypical a – b.” *ELOPE: English Language Overseas Perspectives and Enquiries* 11 (2): 8–20. <https://doi.org/10.4312/elope.11.2.7-20>.
- Goatly, Andrew. 2007. *Washing the Brain: Metaphor and Hidden Ideology*. Amsterdam: John Benjamins.
- Goźdz-Roszkowski, Stanisław. 2011. *Patterns of Linguistic Variation in American Legal English*. Berlin: Peter Lang.
- Goźdz-Roszkowski, Stanisław, and Gianluca Pontrandolfo. 2015. “Legal phraseology today: corpus-based applications across legal languages and genres.” *Fachsprache* 37 (3–4): 130–38. <https://doi.org/10.24989/fs.v37i3-4.1287>.
- Gries, Stefan. 2008. “Phraseology and linguistic theory: A brief survey.” In *Phraseology: An Interdisciplinary Perspective*, edited by Sylvianne Granger and Fanny Meunier, 3–25. Amsterdam: John Benjamins.
- Grimm, Jacob. 1816. “Von der Poesie im Recht.” *Zeitschrift für geschichtliche Rechtswissenschaft* 2: 25–99.
- Herrera-Soler, Honesto, and Michael White, eds. 2012. *Metaphor and Mills: Figurative Language in Business and Economics*. Berlin: de Gruyter.
- Imamović, Adisa. 2013. “Metaphor and metonymy in legal texts.” *Jezikoslovje* 14 (2–3): 295–306.
- Kjær, Anne L. 1990. “Normbetingede ordforbindelser i tysk juridisk sprog.” Phd diss., University of Copenhagen.
- . 2007 “Phrasemes in Legal Texts.” In *Phraseologie/Phraseology. Ein internationales Handbuch zeitgenössischer Forschung/An International Handbook of Contemporary Research*, edited by Harald Burger, Dmitrij Dobrovolskij, Peter Kühn, Niel Noerrick, 506–15. Berlin: de Gruyter.

- Konecny, Christine. 2010. *Kollokationen. Versuch einer semantisch-begrifflichen Annäherung und Klassifizierung anhand italienischer Beispiele*. [Forum Sprachwissenschaften 8]. München: Martin Meidenbauer.
- . 2012. "Wirf' mal einen Blick darauf – Dacci un po' un'occhiata! Zu interlingualen Abweichungen in italienischen und deutschen Verb-Substantiv-Kollokationen und deren Relevanz für die Translationsdidaktik und die translatologische Praxis." In *Translationswissenschaft interdisziplinär: Fragen der Theorie und Didaktik*, edited by Lew Zybatow et al., 299–306. Frankfurt am Main: Peter Lang.
- Kopaczyk, Joana. 2013. *The Legal Language of Scottish Burghs. Standardization and Lexical Bundles 1380–1560*. Oxford: Oxford University Press.
- . 2017. "Terms and conditions: a comparative study of noun binomials in UK and Scottish legislation." In *Phraseology in Legal and Institutional Settings: A Corpus-Based Interdisciplinary Perspective*, edited by Stanisław Goźdz–Roszkowski and Gianluca Pontrandolfo, 160–77. London: Routledge.
- Kordić, Ljubica. 2023. "Metaphors Lawyers Live By." *International Journal for the Semiotics of Law* 36: 1639–54. <https://doi.org/10.1007/s11196-023-09975-0>.
- Kövecses, Zoltan. 2006. *Language, Mind and Culture: A Practical Introduction*. Oxford: Oxford University Press.
- Lakoff, George, Jane Espenson, and Alan Schwartz. 1991. *Master Metaphor List. 2nd Draft Copy*. Berkeley: Cognitive Linguistics Group, University of California at Berkeley. <http://araw.mede.uic.edu/~alansz/metaphor/METAPHORLIST.pdf>.
- Lakoff, George, and Mark Johnson. 1980. *Metaphors We Live By*. Chicago: The University of Chicago Press.
- Nesselhauf, Nadja. 2004. "What are collocations?" In *Phraseological Units: Basic Concepts and their Application*, edited by David J. Allerton, Nadja Nesselhauf, and Paul Skandera, 1–21. Basel: Schwabe.
- Ostroški Anić, Ana. 2019. *Metafora u terminologiji*. Zagreb: Institut za hrvatski jezik i jezikoslovlje.
- Patekar, Jakob. 2022. "What is a metaphorical collocation?" *Fluminensia* 34 (2): 31–49.
- Philip, Gill. 2011. *Colouring Meaning. Collocation and Connotation in Figurative Language*. Amsterdam: John Benjamins.
- Pragglejaz Group. 2007. "MIP: A method for identifying metaphorically used words in discourse." *Metaphor and Symbol* 22: 1–39.
- Reder, Anna. 2006. "Kollokationsforschung und Kollokationsdidaktik." *Linguistik Online* 28 (3): 157–76. <https://doi.org/10.13092/lo.28.617>.
- Richard, Isabelle. 2014. "Metaphors in English for Law: Let Us Keep Them!" *Lexis, Journal in English Lexicology* 8: 1–19. <https://doi.org/10.4000/lexis.251>.
- Sinclair, John M. 2004. *Trust the Text: Language, Corpus and Discourse*. London: Routledge.
- Stojić, Aneta. 2015. "Kollokationen in der Sprach- und Kulturmittlung (am Beispiel des Sprachenpaares Deutsch-Kroatisch)." In *Translation und transkulturelle Kommunikation*, edited by Annette Đurović and Vlasta Kučič, 81–96. Belgrade: University of Belgrade.
- . 2024. *Metaforičke kolokacije – od konceptualizacije do ustaljene sveze*. Rijeka: Filozofski fakultet.
- Stojić, Aneta, and Iva Barić. 2013. "Acquisition of collocations – Goals and methods." In *Neue Methoden beim Fremdspracherwerb, dargestellt am Blended Learning*, edited by Mehmet Uysal, 79–104. Isparta.
- Stojić, Aneta, and Nataša Košuta. 2022. "Izrada inventara metaforičkih kolokacija u hrvatskome jeziku – na primjeru imenice godina." *Fluminensia* 34 (1): 9–29.
- Stojić, Aneta, and Mihaela Matešić. 2024. "Metaphorical potential of lexical collocations: the case of the noun *prilika*." *Slavia Centralis* (forthcoming).
- Stojić, Aneta, and Sanela Murica. 2010. "Kolokacije – teorijska razmatranja i primjena u praksi na primjerima hrvatskoga i njemačkoga jezika." *Fluminensia* 2 (2): 111–25.
- Stojić, Aneta, and Tamara Stiglič. 2011. "Kollokationen im deutsch-kroatischen Sprachvergleich." *Jezikoslovlje* 12 (2): 263–82.
- Volungevičienė, Skaistė. 2008. "Metaphorische Kollokation: Zwischen Metapher und Phraseologismus." *Kalbotyra* 59 (3): 1392–1517.
- Winter, Steven. 2001. *A Clearing in the Forest: Law, Life, and Mind*. Chicago: University of Chicago Press.

Wojtczak, Sylwia, and Ivona Witczak-Plisiecka. 2019. "Metaphors and legal language: A few comments on ordinary, specialised and legal meaning." *Research in Language* 17 (3): 273–95. <https://doi.org/10.18778/1731-7533.17.3.04>.



## Introduction of Machine Translation into Audiovisual Translation Teaching

### ABSTRACT

Subtitling has long been the primary type of audiovisual translation in Slovenia. A steady increase in the production of audiovisual content, including English-speaking movies and TV series, along with the development of machine translation, have seen a growing need and opportunity for subtitling globally. The Department of Translation Studies (University of Ljubljana) has offered courses in subtitling to MA students for over twenty years. In the course of this time, teaching methods have been adapted to embrace new technologies. Recently, students have shown interest in incorporating machine translation into the subtitling course. The paper presents an analysis of a subtitling assignment in which students were asked to post-edit DeepL's Slovene translation of English formatted movie subtitles. The post-edited subtitles were compared with the subtitles done by a group of students from scratch. The results show considerable differences in the students' multimodal awareness and the overall quality of the subtitles between the two groups.

**Keywords:** subtitling, AVT teaching, machine translation, post-editing, multimodal awareness

## Uvajanje strojnega prevajanja v poučevanje podnaslovnega prevajanja

### IZVLEČEK

Podnaslovno prevajanje je primarna zvrst avdiovizualnega prevajanja v Sloveniji. Zaradi razmaha ustvarjanja avdiovizualnih vsebin, ki vključujejo filme in nadaljevanke v angleščini, in razvoja strojnega prevajanja se po vsem svetu večajo potrebe in možnosti za podnaslovno prevajanje. Oddelek za prevajalstvo Univerze v Ljubljani že več kot dvajset let nudi študentom magistrskega študija seminarje iz podnaslavljanja. V tem času so metode poučevanja sledile novim tehnološkim možnostim. Študentje v zadnjih letih kažejo zanimanje za vključevanje strojnega prevajanja v seminar podnaslavljanja. Članek predstavlja analizo naloge iz podnaslavljanja, v kateri so študentje popravili in uredili slovenski prevod angleških filmskih podnapisov, ki ga je ustvaril prevajalnik DeepL. Tako urejene podnapise smo primerjali s podnapisi, ki jih je druga skupina študentov ustvarila v celoti. Izsledki kažejo precejšnje razlike v multimodalnem zavedanju in nasploh kakovosti podnapisov med študenti obeh skupin.

**Ključne besede:** podnaslavljanje, poučevanje avdiovizualnega prevajanja, strojno prevajanje, popravljanje strojnih prevodov, multimodalno zavedanje

# 1 Introduction

New technologies have seen an increase in new audiovisual media services over the last twenty years. According to the Audiovisual Media Services Directive issued by the European Commission in 2010, audiovisual media services are defined as “mass media in their function to inform, entertain and educate the general public”, including related non-private commercial communication, while excluding subsidiary occurrences of audiovisual elements in advertising and gambling activities (European Commission 2010, 3). The Directive highlights the *educational* impact of the audiovisual media on their users. In fact, in the years prior to the adoption of the Directive, spanning the two EU enlargements in 2004 and 2007, the European Commission published several documents in which it stressed the importance of developing multilingualism via language learning methods.

In its 2008 Communication on Multilingualism, the European Commission conveyed its belief that “[t]he media, new technologies and human and automatic translation services can bring the increasing variety of languages and cultures in the EU closer to citizens and provide the means to cross language barriers” (European Commission 2008, 12), stressing that languages can be informally learned via the media. These services “can thus also be a great source of informal language learning through ‘edutainment’ and subtitled films” (European Commission 2008, 12). To promote this objective, the Commission has supported subtitling and invited the Member States to encourage it. Two years later, in 2010, the Commission launched a 12-month project in 33 European countries aimed at exploiting the true potential of subtitling to encourage language learning and strengthen foreign language competence. While the final report waters down the potential of subtitling in this context, it concludes that subtitling helps to “improve the mastery of foreign languages”, can “raise awareness and provide motivation for language learning” in pupils, students and immigrants, with multilingualism and higher education being as such indicative of a preference for subtitling over dubbing (Media Consulting Group 2011, 26). Despite the Commission’s call for linguistic diversity, the study showed the prevalence of North American film production and, consequently, the prevalence of English over other languages. The Commission recommended the use of subtitling as an educational tool in both the expected educational environments, i.e., (secondary) schools, (all) universities and language teaching organizations, as well as in actions by broadcasters. One of the Commission’s strategic recommendations was not only to *use* subtitles to support language teaching and learning, but also to encourage “students and pupils” to *create* subtitles (cf. Media Consulting Group 2011, 28). The underlying ease of this proposal is surprising since it takes a whole MA course to master the theoretical and practical rudiments of subtitling. Such a recommendation may lead to the reinforcement of the stereotype that subtitling does not require any special training while, paradoxically, viewers will, deservedly or not, violently criticize the quality of TV subtitles. Last but not least, the Commission recommended involvement of “media professionals to develop and/or make available quality European films in subtitled versions”, especially in languages of limited scope, as well as to encourage especially young adults in predominantly non-subtitling countries to watch subtitled programmes (Media Consulting Group 2011, 29). The Commission’s recognition of the educational scope of subtitled programmes has been seconded by the rise of new audiovisual media services.

In line with the recommendations above, the European Commission legally regulated the nature and provision of the services with the Audiovisual Media Services Directive in 2010. Directive 2010/13/EU stipulated the accessibility of audiovisual media services to ensure “freedom of information, diversity of opinion and medial pluralism” and that the provisions of programmes should also cover accompanying text-based content, “such as subtitling services and electronic programme guides” (European Commission 2010, 3). The Directive was amended in 2018 to include programmes and user-generated content shared via social media services (European Commission 2018, 69–70), summoning the EU Member States to make audiovisual content immediately accessible especially to persons with visual and hearing impairments, with subtitling being again listed as one of the features to this end.

The European Parliament adopted a report in early 2023 in which it urged the Member States to facilitate accessibility through dubbing, subtitling, audio description or other means in all languages, both official and regional and minority languages (Kammerevert 2023). The onset of neural machine translation and large language models did not go unnoticed, as the 2023 report brings up the “potential of using artificial intelligence” as the implementation tool and “calls on the Commission and the Member States to promote this in a strategic and targeted manner” (Kammerevert 2023, 13). More specifically, artificial intelligence (AI) should be used in the production and distribution of the AV content to promote freedom of expression, the exchange of information and “detect illegal content” (Kammerevert 2023, 29–30). In fact, the European institutions have been using AI to promote their video content, while the EU web portal Europeana has offered high-quality video clips on European heritage to both professionals and the general public to proofread their AI-generated subtitles and captions, notably through a resource developed by the 2021–2022 Europeana Subtitled project (Europeana, n.d.).

## 2 Audiovisual Machine Translation

While research on machine translation (MT) started in the USA in the 1950's, it was only around the turn of the millennium, about ten years after a dramatic rise in academic research on audiovisual translation (AVT), that the technical, rather than semantic, aspects of subtitling drew the attention of MT researchers.<sup>1</sup> Although software developers would present their findings as successful and promising (see Popowich et al. (2000) for the English-Spanish language combination, Armstrong et al. (2006) between German and English), the training material was ill-suited to the context, and the project results regarding semantic accuracy were vague. Nevertheless, the US company Global Translation Inc. still provided live machine translation of audiovisual (AV) content in eight languages, although Díaz Cintas (2005, 20–21) noted the dubious quality of its product, suggesting that the introduction of the tool was an economically driven enterprise.

### 2.1 Statistical Machine Translation in Subtitling

Modern MT has made tremendous leaps in quality since the introduction of the first open-source statistical MT engines in the mid-2000's, most notably the Google Translate service

---

<sup>1</sup> For a detailed survey of MT development of subtitling, see Bywood et al. 2017.

in 2006. The first batch of 30 languages supported by Google Translate was headed by English, which has also functioned as a relay language for most of the GT languages. Slovene was added in 2008, becoming one of the first 40 GT-supported languages. Koletnik Korošec (2011, 16–17) provides useful insight into the response of Slovene students of translation to GT-generated texts, and discusses the didactic potential of MT. Also supported by the findings in Koletnik Korošec (2011), the students were aware of MT's limitations at the time. Unless used in predictable textual environments, early GT-generated texts were derided as a source of comical linguistic transformations, occasionally making the service a *toy* rather than a *tool*. Simultaneously, research on subtitling using statistical MT was developed to incorporate increasing huge training corpora. Once again, the methodology was debateable, as (due to copyright protection) the materials would be collected from freely available subtitle files on the internet, as done by Tiedemann (2007). Sadly, the subtitles were mostly produced by amateur subtitlers with insufficient linguistic and subtitling competences.

Volk (2009) built a statistical MT system for subtitling from English between Swedish and Danish, and had automatically produced subtitles post-edited by professional translators, who showed different tolerances to the quality of the subtitles, and noted that quality varied across film genres. His findings showed that the system was able to achieve good results with a large high-quality parallel corpus, and even though the automatic translation quality was evaluated differently, the customer was convinced of the system's benefits and started to employ the system.

Volk and Harder (2007, [1]) saw the direct written representations of spoken discourse in subtitles as less than ideal, and highlighted the need for textual condensation in subtitling. In the beginning, subtitles were perceived as easily trainable material, since they were regarded as short, syntactically simple, repetitive textual units embedded in comfortably alignable timecodes. Working with corpora has shown that since subtitles feature a wide range of genres and elements of literary language (register, wordplay, poetry) and a variety of topics and styles, movie subtitling is closer to the translation of fiction, especially plays, rather than non-fiction.

The use of statistical MT was explored by the SUMAT (SUBtitling for MACHine Translation) project between 2011 and 2014. The aim of the EU project was to develop a commercial online post-edited subtitling service which would be used with AV programmes broadcast by public TV networks. The system developers pointed out that the expansion of the market was hindered by issues “such as cost, time and quality” (Petukhova et al. 2012, 21; Sepesy Maučec et al. 2012, 167), and aspired to higher productivity for lower economical remuneration. The project involved nine European languages, with translation between English and Dutch, French, German, Portuguese, Spanish, and Swedish, while the Serbian-Slovene pair was chosen as “a test-case of an under-resourced language pair” (Etchegoyhen et al. 2014, 46). The training corpora included aligned parallel subtitles, most of which were collected from the unprofessional OpenSubs website. Despite the good results on objective ratings and an average gain of almost 40% in productivity, the post-editing process was not evaluated positively by professional AV translators, whom the authors referred to as “a key aspect for any eventual adoption of machine translation technology in professional subtitling” (Etchegoyhen et al. 2014, 46).

Even though the service was aimed at professional subtitling, the authors would strive to further improve the quality of machine translation to reduce “the cognitive effort in post-editing machine translated text in the open subtitling domains” (Etchegoyhen et al. 2014, 46, 52).

Providing another view of the SUMAT project, Bywood, Georgakopoulou and Etchegoyhen (2017, 502–3) aim at a more energetic integration of the MT technology into the subtitling workflow. They see the need for the creation of a new profile – “subtitle post-editor” – with either the “existing subtitlers” being trained to meet the needs of post-editing, or training existing post-editors in subtitling skills (Georgakopoulou and Bywood, 2014, 28 in Bywood 2017, 502), “to seek out people whose skills and experience include high levels of attention to detail and possibly a tolerance for repetitive work and train them specifically in subtitle post-editing” and to make postgraduate AVT courses embrace the need for post-editing.<sup>2</sup> This alienating *corrective* rather than inspiring *creative* focus on translation opens a bleak future for the AVT profession.

## 2.2 Neural Machine Translation in Subtitling

The mid-2010’s marked a quantitative leap in machine translation. Artificial neural networks are now capable of predicting the likelihood of syntagmatic and paradigmatic structures in text production. The new approach has introduced new terminology which takes the mechanical imagery to *organic* images: machines have been increasingly thought of as forms of (artificial) *intelligence*, the notion of externally induced training has been replaced by deep *learning*, while service development has been rephrased as AI *evolution*. The human-like interface was further enhanced by the introduction of ChatGPT in late 2022, a generative pre-trained transformer, developed by OpenAI. ChatGPT makes use of deep learning architectures and large learning language models, and is often presented as a chatbox mimicking human conversation.

As of 2017 there have been two open-source neural machine translation services which have gained most popularity: Google Translate and DeepL. The latter is a Germany-based machine translation provider, which added support for a dozen more European languages in early 2021, when Slovene became one of its 26 languages. Despite the Department’s initial discouragement with regard to employing AI services in postgraduate translation classes, the students’ accidental copying of MT sources into their assignments revealed they were using and post-editing the MT texts as best as they could. The recent leap forward in neural machine translation coincided with the COVID-19 pandemic, which upset the regular educational patterns between early 2020 and mid-2021. Other than the deplorable lack of physical in-class experience, which is conducive to the process of human learning, the pandemic did not interfere with the very technical aspects of AVT teaching. Indeed, it may have been the greater reliance on digital technologies during the two years of studying from home that sparked a stronger interest in the use of MT among (technology-minded) students, especially the promising solutions from DeepL.

<sup>2</sup> See Mezeg (2023) for a discussion on the needs of developing post-editing skills in students of translation.

Indeed, DeepL has proven to be semantically more accurate than Google Translate, and recent developments in AI technologies have further improved the generation of automatic or real-time intralingual and interlingual subtitles (Poibeau 2023). The last few years have seen several new platforms that allow automatic (or manual) transcription and translation of video audio, such as Descript, Matesub, Rev, Sonix, Whisper. Online subtitle services claim they provide “professional” video editing software to modify the content and export files in industry standard formats. The services advertise almost a 100% accuracy, even though their platforms are integrated with Google Translate, which has proven inferior to DeepL. The option of (post-)editing by such platforms, however, does not differ greatly from that when using the free software application Subtitle Edit, aimed at subtitle creation according to user-defined time and space restrictions.

The quality of the original English subtitle files that (Slovene) professional AV translators are given as templates has been declining in recent years. The extreme lengths of lines containing the entire utterance, hardly manageable reading speeds, occasionally misheard words and awkward time spotting may indicate that such files have been automatically generated without observing any high-quality subtitling standards. Editing such files to meet the intended subtitling standards can be more time consuming than creating subtitles from the first step.

Digital translation resources used to be seen as tools in the hand of a translator, but now they appear to have been getting out of (their) hand. AI language models have been superseding manual translation, the quality scope however, has been limited to informative, repetitive, preferably (well)-written discourse for a limited number of language pairs. Poibeau (2022, 6018) reassesses machine translation in terms of human intelligence and shows that “human parity” and “super human performance” is possible for a limited number of languages, text types and “with very literal translations”.

Recent studies on automatic subtitling tools presented in the Findings of the IWSLT 2023 Evaluation Campaign show that “automatic solutions do not reach the level of quality that is necessary in subtitling”, and observe that it is TV series that pose greatest challenges for automatic subtitling (Agarwal 2023, 13–14, 15), although such challenges also apply to movies.

Research that shows awareness of professional subtitling standards thus reveals persistent deficiencies in MT systems, most notably in being unable to tackle the multimodal nature of AV texts. If anywhere, it is in high-quality subtitling that generative transformers may or will develop systems to achieve communicatively suitable and readable subtitles sensitive to the relation between speech, image and different levels of contextual meaning.

### 3 Methodology

Ljubljana MA students are invited to use Subtitle Edit (SE),<sup>3</sup> a free and open-source subtitle editor developed by the Danish IT expert Nikolaj Olsson. The programme has a built-in machine translation option with Google Translate V1 API. In addition, Olsson has developed his own machine translation programme from Swedish to Danish, i.e., two languages that are

---

<sup>3</sup> The subtitling software is available at <https://www.nikse.dk/subtitleedit>.

linguistically and culturally related. Even so, Olsson stresses the need for manual correction of the automatic translation. As the quality of Google Translate for translation between English and Slovene does not reach professional standards, the students of subtitling have not been encouraged to use the in-built feature.

During the academic year 2022/23, ten first-year MA students enrolled in my 60-hour seminar in AVT of movies and series and some assignments in other genres, i.e., animated films for children and documentaries. The group consisted of eight native Slovene speakers and two students who had lived outside Slovenia, both speakers of another Slavonic language. Except for the bilingual student with some video fan-subtitling experience, the students had no prior experience in subtitling. Students were first given some assignments through which they gradually built their AVT competence from the beginning.

One third into the seminar they were given a subtitling assignment based on a translation by DeepL Translator.<sup>4</sup> The students were asked to post-edit a four-minute passage from the 2014 romantic comedy *And So It Goes* directed by the American director Rob Reiner, starring Michael Douglas and Diane Keaton. The film was critically panned for the clichéd script that did a disservice to the great performances by the leading actors, while careful commentators criticized the film's rating for foul and inappropriate language, sexuality and substance abuse.

Despite the poor reviews upon the movie's release, the dialogues provide worthwhile subtitling challenges. Prior to the MT home assignment, the students had been working on the film, which would help them to consider appropriate subtitling strategies for the genre of romantic comedy. They had become acquainted with the story's cultural background, the psychological profiles of the main characters, the interpersonal relations (professional and social relations between the characters spanning three generations), as well as situation-specific communication. The students had gathered some experience in observing textual segmentation and adaptation as well as acquired translation strategies that covered stylistic features of register, informal conversation, technical language, child's language, emotional language, idiomaticity, humour, sarcasm, cultural specificity (localization), political (in)correctness, mild profanity and song translation. While these technical, stylistic and pragmatic aspects of subtitling should be crucial to any language pair, Slovene AV translators also have to observe the complexity of Slovene morphology. They may need to express features that in English remain unmarked, such as grammatical indicators of gender, number, and formality. The relevant information may be found in the immediate context or visuals, or elsewhere in the movie.

The students were given English time-coded subtitles containing entire speeches, except for very few minor omissions and graphic texts. They were asked to form a maximum of two-line subtitles with one speaker per line and observe the (currently valid) national TV parameters of a maximum of 35 characters per line and 12 characters per second for two-second+ subtitles to allow for a comfortable reading speed of large numbers of viewers watching the PG-13 movie. They were asked to consider the subtitling strategies that they had been acquiring and improving in previous assignments and to deploy new strategies as necessary. The students were asked to run the text through DeepL Translator, post-edit the translation in accordance

---

<sup>4</sup> The web-based translation service is available at <https://www.deepl.com/translator>.

with the instructions given above and comment on their experience. The results of the 2023 “MT group” (with subtitles provided by a maximum of ten students)<sup>5</sup> were compared with the results of the 2021 “non-MT” group (with subtitles provided by a maximum of 15 students).<sup>6</sup> The “non-MT” students had received the same instruction in AVT and were given the same assignment, with the four-minute passage split into two parts. The students of the “non-MT” group were asked to translate the passage on their own, i.e., without machine translation.

The four-minute passage from *And So It Goes* contains four scenes: (1) the “Paintball” scene, (2) the “Landlord” scene, (3) the “Compassion” scene, and (4) the “Real Estate” scene. Scene 1 features an emphatic young man’s words of consolation to a stray dog. The man’s (rhetorical) question is a cue that opens Scene 2, which pivots around the self-centred real estate agent and owner of the building, Oren Little. The hum of overlapping multiple speakers indulging in small talk becomes clearer as some tenants accuse Oren of rude and selfish behaviour. In Scene 3, Leah’s comment on Oren’s lack of compassion develops into an embittered widower’s slow-paced monologue. Following some phatic exchanges, Scene 4 features a verbal confrontation between three real estate agents, all eloquently defending their positions. Their factual data are laden with real estate jargon and flanked by cutting retorts that combine marked vocabulary and complex figures of speech.

## 4 Analysis

The analysis studies the differences in translation decision-making between the MT-group and the non-MT group for the grammatical (examples 1–7), lexical (example 8) and stylistic categories (examples 9–12) which show the students’ mental processing in relation to the multimodally derived meaning. As Figure 1 shows, the percentages for all the categories that suggest some affinity with DeepL’s translation strategies are mostly considerably higher with the MT-group students than those with the non-MT-group students.

For the sake of illustrative comparison, the analysis includes comments on automatic translations by two automatic subtitling tools, Sonix<sup>7</sup> and Matesub,<sup>8</sup> added during the time of writing the paper.

### 4.1 English Grammatical Polysemy

Based on or inspired by natural conversation, movie discourse, very much like drama, commonly takes the form of dialogue between the participants in the scene. As conversation generally turns on personal and interpersonal topics between the speakers, the discourse is marked by first-person and second-person references. The English pronominal system is syncretic and thus polysemic. Slovene, on the other hand, uses an extensive system of discrete grammatical

---

<sup>5</sup> One student, a non-native speaker of Slovene, failed to submit some subtitles.

<sup>6</sup> The assignment was then divided into two parts. The first non-MT group consisted of twelve students, the second non-MT group of five students, of whom one bilingual student failed to submit some subtitles.

<sup>7</sup> Sonix’s home page advertises their tool as “the best automated transcription software in 2024” and “the world’s most advanced automated transcription, translation, and subtitling platform” (Sonix, n.d.). The platform is available at <https://sonix.ai>.

<sup>8</sup> The web-based subtitling tool is available at <https://matesub.com>. See Karakanta et al. 2022 for user feedback by subtitlers on Matesub.

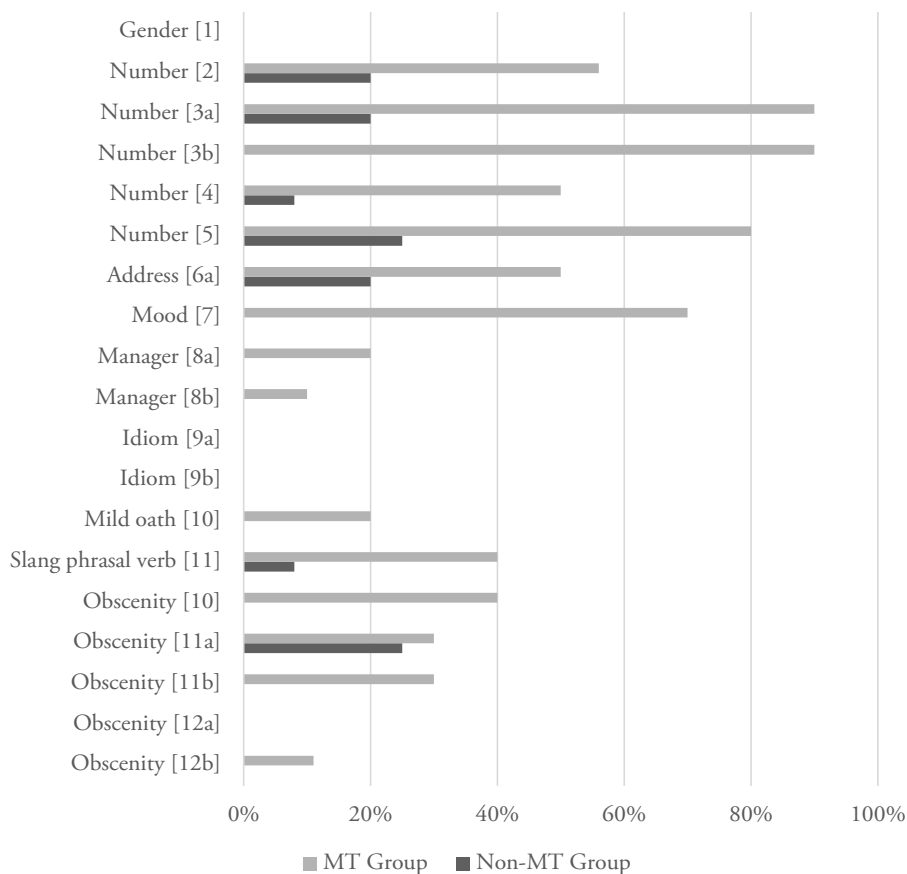


FIGURE 1. DeepL-like (mis)translation strategy: the ratios of DeepL-like translations for the numbered illustrative examples between the MT-group students and non-MT students.

morphemes, such as pronominal and verbal inflexions that indicate the grammatical person for their gender, number and level of formality, while the predicator is further marked for tense, aspect and mood. As the Slovene pronominal-verbal system tends to be monosemic, the translator's interpretation relies on contextual linguistic and extra-linguistic indicators. Poibeau (2022, 6020) states that most current MT systems operate at the sentence level. This paper shows that DeepL Translator does indeed have difficulties interpreting the context beyond the sentence limit, even beyond the line limit, which results in wrong interpretations. Although pronominal anaphora is a prominent discoursal feature in movie dialogues, interpretational problems with MT from English into Slovene may appear at the supra-sentence level in other types of discourse, see Mohar, Orthaber and Onič (2020, 134–35) for examples in the translation of poetic fiction by Google Translate and Translator.eu.

#### 4.1.1 The Subject's Gender

Pronominally, English reveals the Subject's grammatical gender only with third-person references. Having informative finite verb inflections, Slovene omits the Subject pronouns in

unmarked communicative positions. The gender, however, is automatically revealed when the Subject is bound with predicative adjectival or participial grammatical features. Slovene past participles form periphrastic tenses, which, dependent on the narrative, are frequent in movies.

Whereas Poibeau (2022, 6020) finds that current MT systems choose one gender randomly, and are “thus often wrong”, this study shows that with non-third person singular references DeepL-generated translations, done via English, have shown a bias for the *masculine* grammatical gender.<sup>9</sup> Although there are three female characters that present some speech in the subtitled passage there are only masculine Subject references outside the present-tense.

However, it may sound surprising that DeepL chose to translate some first-person sentences in the “Compassion” scene as if spoken by a female I-character. The MT translation may have been affected by the emotional content that is stereotypically attributed to women. The feminine gender was chosen even with two subsequent sentences in the same line. The Slovene translation of Oren Little’s little soliloquy is indicated by gender symbols following the English predictors:

[1] (Oren Little:)

My wife died of cancer!

For two years, I **bathed**<sup>10</sup> (M) her, I **changed** (M) her.

I **cooked** (M) for her. I **fed** (F) her.

I even **prayed** (F) for her.

And I **cried** (F) all I’**m** ever **gonna cry** (F).

I **had** (F) compassion.

All the MT students that provided the subtitles for this scene post-edited the feminine participles or rephrased DeepL’s translations accordingly. Although this type of post-editing is not difficult, Slovene AV translators have to maintain a constant gender-oriented focus, while in non-automatic translation the attention can be delegated to other aspects of subtitling. Although automatic subtitling tools have been trained to account for the concept of Slovene grammatical gender, they may fail to maintain the focus beyond the sentence limit. It is interesting to observe that while Sonix’s automatic speech recognition discerns different speakers, the versions of Oren Little’s automatically translated monologue by both Sonix and Matesub reflect similar gender-specific emotional stereotypes as seen with DeepL. The gender-inflexion issue was not relevant to non-MT students, as there were sufficient multimodal sex/gender indicators that are easily spotted by human translators.

#### 4.1.2 Grammatical Number

Grammatical number is a linguistic feature morphologically assigned to nouns, pronouns in both English and Slovene, as well as adjectives and participles in Slovene. While English

---

<sup>9</sup> The masculine gender bias may not be limited to the grammatical case: Mohar et al. (2020, 134) observes the gender-related mistranslation in the Direct Object.

<sup>10</sup> All the relevant features in the subtitles are highlighted by Orel Kos.

is marked for singular and plural numbers, Slovene is the only official language among the EU member states to distinguish the dual, a category marking two entities. As movie narratives often address pairs of characters performing joint actions, AV translators need to observe when an English plural will imply two entities. The interpretation may be indicated or supported through the immediate image and sound, or derived from other multimodal information elsewhere in the movie.

The subtitled passage contains three utterances with the pronoun *we* that neither DeepL, Sonix nor Matesub interpreted as containing the dual. The “Real Estate” scene contains an utterance expressed by Claire, an aging real estate agent siding with her retiring colleague Oren Little against her young grandson when the three of them meet in the office:

[2] (Grandma:)

We’ve been here since your grandfather started this business 44 years ago.

In the non-MT group only one (20%)<sup>11</sup> (i.e., bilingual) student out of five students used the plural in Slovene. In the DeepL group five (56%) out of nine students retained the MT plural reference and four adapted it to dual. The ratio of interpretations with the wrong number is considerably higher with the DeepL group.

Interpretation of the “intended” grammatical number is made difficult when there is a mere pronoun-predicator agreement against the visuals with several entities, which happens in the “Landlord” scene. The first *we* reference comes from an African-American man conversing with other neighbours:

[3a] (Tenant:)

You know if he parked over just a little bit, **we** wouldn’t have to hunt for a space on the street.

At this stage in the scene, the number suggested by the pronoun *we* may be perceived as a collective plural or open to speculation. A minute later, a foregrounded dialogue between the man and Oren Little offers a clearer interpretation:

[3b] (Tenant:)

My wife is going to give birth soon. Yesterday, **we** had to park a block away because you refuse to move your car just a few feet.

The dual concept in the pronoun *we* is hinted at by the man’s reference to his pregnant wife, while the rest of the neighbours sit listening in the background. None of the ten MT students post-edited the plural references in (a) or (b) to express the dual. As many as eight (80%) students retained DeepL’s plural references. Two students partially adapted the grammatical number, and failed to interconnect the two subscenes (a) and (b), applying a plural (a) + dual (b) and a less expected dual (a) + plural (b) combination respectively. In contrast, ten out of

<sup>11</sup> The percentages are given for the number of the non-MT students that provide DeepL-like translations and the MT-students that retain DeepL’s solutions.

twelve non-MT students applied dual to both utterances, while two non-MT students chose the partially contextualized switch from plural (a) to dual (b). No non-MT students rendered the English plural references as entirely plural in Slovene.

All students, the MT group included, had been instructed to check or pay attention to English plural references against the visuals and discourse or elsewhere in the movie. In an assignment preceding the MT task discussed in the paper, a scene with the visuals of some children included a similar plural/dual to dual combination of dialogues. The “MT” group students, who then translated from scratch, achieved more contextualized translations: five (50%) students out of ten fully recognized the dual references, while the “non-MT” group achieved even better results: eleven (78.5%) out fourteen students interpreted the English plural references as dual in Slovene.

### 4.1.3 Levels of Formality

Modern Slovene has three levels of address: formal address, informal address, and semi-formal address. The forms reflect vertical and/or horizontal relationships to a singular addressee. The formal v. informal opposition is the default system for both human dominance and distance. The grammatical features are marked pronominally and verbally. The formal address uses the second person plural and, where syntactically required, masculine, i.e., linguistically unmarked generic gender inflections. The semi-formal address is a combination of formal and informal grammatical features, it takes the second person plural finite forms and singular gender inflections for participles and adjectives. The semi-formal address regulates public and social distance, for example in customer services to imbue a formal conversation with some emotional proximity. It is used colloquially and is not normatively recognized. Unless presented as a character’s overt stylistic feature, this grammatical mixture is not tolerated in professional Slovene subtitles. Compliant with the genre and the character’s sociolinguistic profile, Slovene professional subtitles may use substandard vocabulary but no substandard grammatical features. Grammatically, all Slovene levels of address to singular referents are covered by the English pronoun *you*. The pronoun’s syncretic nature also accounts for Slovene dual and plural and their respective gender markers as well as non-Subject functions.

English has no grammatical markers for the formal-informal opposition. Attitudes of deference are readily expressed lexically and allow for greater swings or immediate transitions in familiarity. The speakers’ swings in emotional proximity can be reflected in subtitles of inexperienced or incautious AV translators into Slovene who will switch between formal and informal addresses to accommodate the implied decrease or increase in social distance.

AV dialogues teem with polysemic *you* references, including those implied by the Imperative mood. The Slovene AV translator derives their (own) optimal multimodal interpretation from the number of people involved, social situations and the characters’ personal traits. The assignment’s automatic translations show that multimodally blind MT tools (DeepL, Sonix, Matesub) heavily rely on the lexical information in the clause when to choose between the second person singular and the second person plural with (generic) masculine markers. Firstly, basic conversation and simple familiar exchanges with *you* references, e.g., *Are you OK? Hold up the towel!*, are assigned the second person singular, suggesting that utterances are

aimed at an individual person. The register sensitive DeepL appears to employ the strategy systematically, Sonix shows a low informal-to-formal threshold, while Matesub seems to follow a similar lexis-based strategy, but there are substantial fluctuations. Secondly, utterances with *you* references that contain substandard or vulgar vocabulary correlate with a discrete person and a lack of human distance. The three automatic tools used the informal second person singular in Slovene to render the utterances that contain substandard vocabulary: *Now what dumbass would shoot you with a paintball gun?*, *Something to think about next time you **get an itch to rat me out**?*, *Don't be a smart-ass*. Thirdly, utterances with *you* references that contain professional terminology and vocabulary that hints at a non-informal conversation correlate with the Slovene second person plural, which may either be interpreted as the formal address or an actual plural reference. The three automatic tools used the second person plural to render in Slovene the utterances that contain some possibly business jargon: *Complain to the manager*, *I don't care if you **do own the building***, *How long have you **been in the business**?*, *You'll **get five-eight, not a penny more***.

The three lexically-based strategies appear to have been systematically used by DeepL, producing the same translations in late 2022 and late 2023, when I had the text retranslated. DeepL would thus refrain from directing substandard vocabulary at several speakers or from using the formal address, while utterances with professional terminology are unlikely to be attributed to individuals with familiar interpersonal relationships.

The “Landlord” scene shows a conversation between Oren Little, a grumpy tenant who turns out to be the landlord, and a family man, who either speaks for himself or represents a cohort of dissatisfied tenants. Over a stretch of 40 seconds the following utterances by the landlord are directed at the family man:

[4] (Oren Little:)

**You called** the manager **on me**. (Slovene: 2nd sg)<sup>12</sup>

[5] (Oren Little:)

Something to think about next time **you get an itch to rat me out**? (Slovene: 2nd sg)

Nine out of twelve non-MT students used the second person plural with both utterances, interpreting the *you* references as either the formal address to refer to the family man or the second person plural to refer to the entire body of complaining tenants. Two non-MT students used a plural-to-singular switch and only one student (8%) used the informal singular address with both utterances suggesting that the old man is presented as a rude person. This personality trait can be supported by the man's use of vulgar language and general attitude towards the young tenants, who are indignant at his inconsiderate parking habits. As translated by DeepL, both utterances used second person singular references in Slovene. Disregarding the fact that the narrative is open to several interpretations, five (50%) out of ten MT students followed the informal singular pattern provided by DeepL, three MT students used a plural-to-singular switch, and two MT students chose the second person plural with

<sup>12</sup> The brackets contain Slovene forms as suggested by DeepL.

both utterances. While the optimal interpretation is open to discussion, the results show that the interpersonal relationships were presented in radically different ways by the two groups of students. The comparable translations by Sonix and Matesub show the plural-to-singular switch. The second person plural in the first utterance probably derives from its rather neutral lexical features in comparison with the informal ones in the second.

The “Real Estate” scene contains a mini dialogue between the grandmother and her grandson, both working in the family real estate business. Although in Slovene extended family relations the formal address is (still) observed by younger generations speaking to older generations, the opposite direction has never been the case. In DeepL’s Slovene translation the grandmother and her grandson maintain a distanced relationship:

[6]

[6a] (Grandma:)

– How long **have you been** in the business? (Slovene: 2nd pl)<sup>13</sup>

(Grandson:)

– Seven years, as **you** well **know**. (Slovene: 2nd pl)

(Grandson:)

– Every day is better than the last.

[6b] (Grandma:)

– **Don’t be** a smart-ass. (Slovene: 2nd sg)

The non-MT group consisted of five students to subtitle this scene. Four students interpreted the grandmother’s statements informally, while one (bilingual) student (20%) used the DeepL-like switch from formal to informal register. Five MT students (including the bilingual and the non-native speaker) (50%) out of ten students chose the informal address. Three students took up DeepL’s formal-to-informal switch or may have forgotten to post-edit the MT formal interpretation in the first statement: a subtitle in the continuation shows that they were aware of the grandmother’s informal, patronizing attitude towards her grandson. Another two students disregarded the interpersonal relationship between the relatives and even hypercorrected DeepL’s interpretation to contain the formal address in all statements by the grandmother to her grandson.

The “Real Estate” scene results suggest that DeepL’s translation may diminish the AV translator’s attention to contextual meaning or multimodal awareness. A comparison of the results for both scenes shows that it was the same five students who disregarded the context in the “Real Estate” scene that would also follow DeepL’s terms of address in the “Landlord” scene and elsewhere in the assignment. This suggests the students’ general tendency to rely on DeepL’s automatic grammaticalization of the interpersonal relationships.

Even though the translation was barely coherent in places, Matesub’s automatic translation observed the same strategy as DeepL, while Sonix, having a low informal-to-formal threshold, used the formal address with all three relevant utterances.

---

<sup>13</sup> The brackets contain Slovene forms as suggested by DeepL.

## 4.2 Lexis and Style

MT systems have shown a tendency towards direct translation. Dependent on contextual interpretation, there are very few fully matching AV translation units between English and Slovene. This study shows that students using DeepL may be lulled into disregarding both lexico-grammatical and macrocontextual acceptability. This is indicated by their insufficient post-editing. The lexico-grammatical acceptability of the MT translations is lower than with the non-MT students, even though both groups had been receiving the same assignments and comparable teaching instructions. The differences may be generational – also in the sense that each new generation of students is more inclined to rely on MT. The following example will show the two groups' different degrees of tolerance to certain lexico-grammatical features.

At the beginning of the “Compassion” scene, after a heated conversation between the landlord and his tenants is over, Leah, a widowed tenant who has kept quiet so far, rebukes the self-centred landlord:

[7] (Leah:)

You need to have a little compassion. Okay?

DeepL translated the utterance literally: *Morate imeti malo sočutja. Dobro?* Both groups had been previously shown that instructions and commands containing English modal verbs such as *need to* may be paraphrased as Slovene imperative sentences to achieve a wording that is shorter and feels pragmatically more natural. Five out of six non-MT students used the imperative in their translations, while one student used a pragmatically smoother modal structure: *Lahko bi imel malo sočutja*. (Literally backtranslated to English: *You could have a little compassion*.) In the MT-group, seven (70%) out of ten students retained DeepL's translation with the modal verb *morati* (English “must, have to, need to”), two students used the milder modal structure *lahko bi*, whereas only one student used an imperative sentence. In addition to this lexico-grammatical feature, the original utterance contains the weak lexical collocation *have compassion*, which does have a formal equivalent in Slovene but will feel stylistically and pragmatically more adequate if paraphrased. Three out of six non-MT students rephrased the English collocation. The noun *compassion* was replaced by the semantically corresponding Slovene predicative adjective, i.e., *biti (vsaj) malo sočuten* or the English delexical verb *have* was replaced by the semantically stronger verb *pokazati* (English “show”). A paraphrase with the predicative adjective was used by only one out ten MT students.

Further, the assignment shows MT problems with lexical polysemy. Humans examine multiple readings by analysing several threads of multimodal clues, which have so far been inaccessible to MT generators. The problem of resolving a polysemic lexeme will be illustrated by the term *manager*. When the word *manager* takes the full lexicographic form, MT can recognize it as a technical term. In movie conversation, technical terms may occur sporadically and will appear in their short forms to indicate shared knowledge among the speakers, for example “building manager” (Slovene “upravnik stavbe”) becomes *manager* (Slovene “upravnik”) “the listing price” (Slovene “oglaševana cena”) is reduced to *the listing* (Slovene “cenitev, cena”). The reference to “*the manager*” had been used twice in the previous assignment so the students should have been familiar with the exact meaning of the term and the person's gender, i.e.,

“(female) building manager”. DeepL’s translation of the three occurrences in the “Landlord” scene provided two translations: [8a] *vodja* (English literal backtranslation: “leader”), which is used for both the masculine and feminine genders, and the masculine noun [8b] *menedžer*. All twelve non-MT students remembered the previous assignment and chose the appropriate short feminine form *upravnica* or its variant *upraviteljica*. Even though the MT-group had received the same explanation in their previous assignment, only three out of ten students used the appropriately gendered term *upravnica*. One used the masculine form and replaced the feminine pronoun in one of the following subtitles by the masculine to adapt the executive function to the traditional social role. Another student chose the masculine form *menedžer* with all the references in the scene, but missed that the feminine pronoun referred back to the masculine *menedžer*. One student interpreted the term as *lastnica* (English “(female) owner”), two students (20%) used DeepL’s first suggestion *vodja*, while two students failed to see that the three nouns and one pronoun had the same referent: one student combined two masculine and two feminine references, while the second student used the term *uprava* (English “management”) with one of Oren’s two identical instructions: *Complain to the manager*.

The example shows a huge discrepancy between the two groups. The non-MT group considered both their previous translation experience and the immediate context, including the gender clue in the following subtitle that ascertained the information that had been shared with them during the previous assignment. Probably relying on DeepL’s ready-made translation, most students in the MT-group paid little attention not only to the context – which spanned a mere 45 seconds – but also to their previous translation experience.

#### 4.2.1 Stylistic Markedness

Automatic translation has been evolving translation strategies to the extent that rather than presenting a mere word-for-word translation, it may produce an incorrect or hallucinatory translation. In terms of human subtitling, this abortive translation strategy compares to Gottlieb’s (1992, 166) AVT notion of *resignation*, which results in “differing expression, distorted content” and is used with so-called “untranslatable” items.<sup>14</sup>

The “Real Estate” scene contains the English idiom *get the gall (to do something)* and its variation (*not*) *have gall* in two subtitles:

[9]

[9a] (Grandma:)

So, where do you **get the gall** to second-guess the listing of an Oren Little?

[9b] (Grandson:)

I **don’t have gall**, Grandma. I have comps.

The idiom expresses negative evaluation of the addressee’s insolent behaviour (Smith 2023).

<sup>14</sup> Compared to 2022, Google Translate has barely improved its *resignation*-like translation of idiomatic and stylistically marked expressions (see Simič 2022 for a list of some (informal) Slovene figurative phrases compiled by Lothar Orel and their hallucinatory English translations by Google Translate in June 2022). DeepL shows more sensitivity to figurative language, while ChatGPT 3.5 has been returning impressive results for the examples found on Orel’s list.

DeepL recognized the institutionalized figurative expression in the first subtitle. Its translation was semantically correct, yet non-idiomatic: *Od kod vam **toliko drznosti**, da* (English functional backtranslation: “where do you get the audacity to”). The variant *not have gall* in the second subtitle was translated literally: *Nimam žolča, babica*. The Slovene suggests either the grandson’s medical condition of having “no bile” or a quaint sounding metaphorical phrase indication that the speaker “has no indignation”. The Sonix automatic translation rendered both idioms correctly, but non-idiomatically. Matesub transcribed the first “gall” as “goal” and translated it accordingly, whereas the second idiom was rendered with the Slovene idiom *imeti jajca*, which corresponds to the English idiom *have the balls*. The idiom features the opposite polarity, praising the doer’s courage. This idiom was used by one of the MT students, while another student expressed the idea of the doer’s “courage” non-idiomatically. As the two subtitles express a highly informatively relevant and stylistically rich message within seven seconds, one of the MT students reduced the idiom to the (rhetorical) question “why”. The rest of the students, MT and non-MT, provided Slovene non-idiomatic verbs and phrases to express the idea of audacity. While the idiom *have (the) gall* is used to criticize the doer’s action, the idiom *(not) have (the) balls* showcases the courage needed to perform a risky action. It is worth noting that the crudeness of the abusive idiom debases the old lady, who, however, rarely keeps quiet.

#### 4.2.2 Profanity

Offensive and taboo language has been used in movies and TV series to underscore characters’ real-life spontaneity, their emotional responses, personality traits or social background. The greater tolerance for (or perhaps insensitivity to) the use of vulgar language in film discourse over the last few decades has seen increased academic research on related translational and sociolinguistic aspects, such as categorization, translation strategies, medium restrictions and censorship (for example Mattsson 2006, Díaz Cintas and Remael 2007, Trupej 2015, Díaz-Pérez 2020).

The four-minute passage contains several examples of mild to strong expressions of offensive language. Seeing a boy changing his wet bathing suit in the “Landlord” scene, Oren Little cries out:

[10] (Oren Little:)

Oh **for Pete’s sake**, will you cover your dick? I’m trying to eat a sandwich.

Oren uses the euphemistic oath *for Pete’s sake* in the movie three times, suggesting that this old-fashioned expletive is his standard emotional response. The students’ previous assignment contained a scene in which Oren yells out the phrase for the first time. The functions of expletives and some subtitling strategies had been discussed in the seminar, while neither the non-MT nor MT groups had been given a comprehensive presentation of the topic in the lecture format by that point.

Sonix and Matesub are not sensitive to vulgar language and struggled with the speech recognition of the mild expletive, both heard it as “I’ll repeat six”. DeepL rendered the phrase adequately as a rather old-fashioned and mild blasphemous oath *za božjo voljo* (English:

“for the will of God”). Oren uses the explicative frequently enough to regard the phrase an interesting stylistic feature. Still, it is deemed less important than the proposition if the temporal or spatial constraints do not allow for more information. Four out of the twelve non-MT students and two out of the ten MT students omitted the phrase. Eight non-MT and six MT students used short mild oaths in Slovene, while two MT-students (20%), who had shown an affinity with the DeepL translation elsewhere in the assignment, retained DeepL’s translation. None of the non-MT students came up with DeepL’s marked rendering, which is probably little known among the younger generation.

As shown earlier by the *you* references, DeepL appears to derive a grammatical interpretation from the lexis in the immediate context. On the other hand, DeepL’s translation in the assignment indicates that it may find it difficult to provide a stylistically adequate translation for these very lexical items. For example, the American English slang phrasal verb in Oren’s remark:

[11] (Oren Little:)

Something to think about next time you get an itch **to rat me out**.

was rendered with the stylistically neutral and semantically less appropriate verb *izdati* (English: “betray”), also observed in the Sonix in Matesub automatic translations. The verb “betray” was used by only one (8%) out of twelve non-MT students, while four (40%) out of ten MT students retained DeepL’s translation. One MT student omitted the verb altogether, whereas the rest of the MT group, five students, and six non-MT students chose the verb (*za*)*tožiti* (English: “tell on”). It was only the non-MT students, five out of twelve, who thought of the stylistically equivalent slang Slovene verb (*za*)*špecati*. This suggests that machine translation both imposes its translation on the AV student and restricts their quest for a more suitable/creative translation.

DeepL’s seemingly fumbling ways with the taboo word *dick* will be further manifested as a conspicuous intrusion into the student’s decision-making process. The reference in Oren’s utterance [10] *will you cover your dick?* received in DeepL’s Slovene translation the medical equivalent *penis*, whereas in the next set of subtitles DeepL applied new strategies:

[11]

[11a] (Oren Little:)

So I got a **dick** in my face, right? Somehow I’m to blame?

(Sl: *kurac*, En: “dick/cock”)

(Young child:)

He said “**dick**”.

(Sl: ”*kurac*”, En: “dick/cock”)

[11b] (Boy:)

Sorry about my **dick**.

(Sl/En: *penis*)

At first sight, DeepL appears to be torn between the medical and vulgar descriptions. On a closer reading, a logic emerges from the contextual lexis. The predicates *cover one's dick* and *be sorry about one's dick* suggest prudent, self-effacing behaviour, hence the *penis* in translation, while *get a dick in one's face* and the quotation marks correlate with a brazen attitude expressed by the bad word. Sonix's translations were partly influenced by audio mistakes, and backtranslated into English the phrases read: *recover your penis*, *I have a penis in the face*, *he's a jerk*, *I'm sorry about my dick/cock*. Matesub demonstrates no tolerance for vulgar expressions and transfers all the references onomastically, as Dick.

In the four utterances in [10] and [11], three out of twelve non-MT students used the most offensive Slovene expression for the male organ *kurac/kurec*, one student chose the offensive expression *klinec*, which neither reads nor functions well in the utterances, while eight out of twelve students chose the multi-purpose word *tič*, which occupies a central position along the axis of stylistically marked expressions denoting the male organ in Slovene. In the four utterances in [10] and [11], it was only four out of ten MT students who used the word *tič*, three students used the medical term *penis*, two students used the most offensive *kurac*. Not paying due attention to the intertextual relations or relying on DeepL's translation, one student first [10] used *penis*, and the most offensive expressive with the next three utterances [11]. Although no students' renderings correspond entirely to DeepL's translations, the percentages of DeepL-like equivalents are 60% for the MT-group to 25% for the non-MT group.

As the next example shows, DeepL analyses lexical clues for slang and euphemistic words alike. In the "Real Estate" scene, the grandmother and Oren Little, having had 44 years of experience in the business, belittle Ted's professional work ethic by reducing him to a baby. While the three are talking shop, Claire delivers the following spiteful remark about her grandson:

[12]

[12a] (Grandma:)

When he was a baby, he used to play with his **willy** like it was a rambunctious puppy.

(Backtr. from Slovene: "When he was an infant/baby, he used to play with his **watchman** like it was a rambunctious puppy.")

[12b] (Oren Little:)

And if he can help me, I'll play with his **willy**.

(Backtr. from Slovene: "And if he can help me, I'll play with his **dick/cock**.")

It is the immediate lexical context that DeepL searches in order to find a clue regarding the polysemic word *willy*, which in this context is used as a (chiefly British) euphemism for "penis". The Slovene translation for the second reference in *play with (one's)* may result from a higher frequency of written data for the Slovene bad word in combination with the verb phrase "igrati se s (svojim)". The Slovene *čuvaj* (English "watchman") for *willy* appears to be a product of an AI hallucination, bringing together perhaps "a rambunctious puppy", "willy wagtail" and "watchdog". For Sonix and Matesub the two utterances are even harder nuts to

crack. Matesub again prefers male names and translates the phrases as *play with his Willy* and *play with this Woody* [sic]. Sonix uses a male name in Oren's remark *play with this Willy*, while the grandmother's punchline, bizarre as her utterance is even without this translation, becomes *play with his clit* (Slovene *ščegetavček*).

In [12a], all the non-MT students and six out of nine MT students used a child's word for "willy" (Slovene: *lulček*, *lulek*, English "peeny"). Two MT students used the word *penis*, while one student used the most offensive term. No student left DeepL's semantic mistranslation unedited.

Oren's comment in [12b] takes a mere three seconds, so two out of the four non-MT students and five out of the nine MT students resorted to omission and condensation of the phrase *with his willy*, losing the appropriate reference. The textually reduced translations suggested that Oren was either going to play or that he was going to play with (his own) willy or that he would play with "him/it", while the "him/it" reference is not clear enough to make an impression on the Slovene viewer. Two non-MT students and three MT-students used the child's word, while one MT student retained DeepL's offensive term.

## 5 Feedback Analysis

I collected the MT students' feedback in a nonstructured analysis supported by the following guidelines: the students were asked to comment on the time spent on the assignment, the positive and negative points of using DeepL and to provide their general view of MT in subtitling. Nine students provided some feedback. Compared to the time spent on previous assignments, the majority found the MT assignment more time-consuming. Four students spent between three and four hours working on what ultimately resulted in 51 to 57 subtitles. Two students suggested that the assignment took "more time than usual". One student indicated that the task was performed "faster", and produced 52 subtitles, while one student spent two hours for 57 subtitles and another a mere 40 minutes for 64 subtitles. These last two students also appreciated that the overall translation was obtained faster. On the plus side, three students found the automatically generated translation a useful starting point they could work with. Two students thought that "some segments" were well translated. There were, however, more negative responses, and as many as seven out of nine students observed that they had to correct the many MT-generated mistakes. Six students pointed out that substantial textual reduction had to be done. Four students noticed the tool's inability to cope with context-bound translation, and that "humans" are needed to resolve the issues of gender, formality, even terminology. Three students commented on the tool's reliance on the word-for-word translation strategy, and the related uselessness with culture-specific items and idiomatic expressions. Three students also thought that the automatic translation and bad translations made their subtitling more difficult. One student appreciated the timecodes that needed "only some adjustment", and yet this same student was among three to criticize the inaccuracies in the time stamps. With regard to the usefulness of MT, three students appreciated the very tool and found it useful for the translation of difficult to understand phrases, while three students would prefer to produce their own translations. Two students, one from the former and one from the latter group of three, found the experiment an

interesting experience. With regard to subtitling, two students might use MT as a source of inspiration and another two students said it might be suitable for slow-paced movies or movies with “easier content”.

The comparison between the feedback and analysis shows that the students who worked faster using DeepL's translation and original formatting retained more of DeepL's mistakes or introduced their own, since they did not pay enough attention to the multimodal structure of meaning.

Even though digital translation tools are meant to speed up the process of translation and to provide accurately communicated messages, this was not the case in my teaching of AV translation. The time and intellectual concentration devoted to post-editing appears to be counter-productive, while the lack of significant post-editing results in poor quality AV translation.

## 6 Conclusion

The analysis of the subtitles done by the non-MT and MT groups shows considerable differences in the multimodal awareness, which is the basis of any adequate AV content interpretation. The non-MT group students were better at applying multimodal translation skills acquired by that point. The subtitles of the MT group show that the students tended to lose focus on the progressing relation between the text and the wider context, including the image. The MT-students' lack of this integrative activity is indicative of machine translation logic. Poibeau notes that human translators understand the text by establishing intratextual cohesion and “discourse external coherence”, i.e., correspondences between textual elements and the extralinguistic reality: “While the machine can reasonably be expected to handle issues of text cohesion at some point, discursive coherence is out of its reach, since it requires knowledge of the world” (Poibeau 2022, 6020).

As DeepL is not a specific AVT system, the paper studies the categories that to the greatest extent possible lie outside the scope of textual adaptation, which is required to achieve readable subtitles. All the categories presented in the paper were more adequately solved by the non-MT group. The subtitles by the MT group show that most interpretational deficiencies emerge from DeepL-generated translations. DeepL fails to consider the context which lies beyond the sentence level and, as the analysis shows, beyond the line break. Further, the analysis suggests that DeepL (and to different extents both Sonix and Matesub) rely on intrasentential lexis-based algorithms that may prefer certain grammatical features in a polysemic language like Slovene. While DeepL appears to have an overall tendency to opt for the masculine (first-person) Subject, certain propositions are stereotypically or sentimentally associated with the feminine grammatical gender. Furthermore, intrasentential lexis governs the level of formality and the grammatical number of the Slovene finite verb phrase, i.e., substandard vocabulary tends to be associated with informal singular verb features, whereas technical jargon will be accompanied by the formal address. Moreover, lexis in the immediate context may influence paradigmatic stylistic choices, which may lead to either over- or under-estimation of profanity.

While there may have been an overall gap in the subtitling competence between the 2021 non-MT group and the 2022 MT-group, it is important to note that DeepL's automated translation was not successful in bridging any discrepancies between the generations. In this context, questions are raised about the ease of certain EU crowdsourcing initiatives which appeal to the general public to either produce subtitles or proofread their automatically generated subtitles.

As the feedback shows, most students realized that post-editing the subtitles did not take any less time than doing all the work themselves would have. On the other hand, spending any less time on post-editing would mean poorer quality subtitles. The students thus welcomed the possibility of access to MT tools, but were aware of the pitfalls and the time-consuming post-editing process. The study shows that for didactic reasons it is counter-productive to introduce MT at the early stages of teaching AV translation. Just as subtitling skills need to be acquired, post-editing, if required, should be regarded as a skill that can and should be developed. At the very least, students need to learn post-editing as a separate skill to be able to critically assess the quality of an automatically translated text.

MT – or more specifically, automatic subtitling – can only be a safe tool in the hands of professional AV translators who are capable of producing high-quality translations by themselves, while MT software would help them materialize the text at a faster pace. To help achieve this goal, Audiovisual Translators Europe (AVTE) published the Machine Translation Manifesto, in which they advocate for the concept of the “augmented translator”, which places an active human translator at the centre (rather than the end) of text production while using AI-tools “to augment their skills” (Deryagin, Pošta, and Landes 2021, 4).

Post-editing machine-translated texts requires a high level of concentration, as some machine-translated texts may appear to be formally impeccable and readable, while concealing semantic blunders. There is also the danger of impoverishing the store of language by reinforcing repetitive linguistic patterns that may sound acceptable or have been uncritically incorporated. As this study shows, MT can be misleading, post-editing can be time-consuming, and the result can be less remarkable and less rewarding than creative (audiovisual) translation.

## References

- Agrawal, Milind, Sweta Agrawal, Antonios Anastasopoulos, Luisa Bentivogli, Ondřej Bojar, Claudia Borg, Marine Carpuat, et al. 2023. “Findings of the IWSLT 2023 evaluation campaign.” In *Proceedings of the 20th International Conference on Spoken Language Translation (IWSLT 2023)*, edited by Elizabeth Salesky, Marcello Federico and Marine Carpuat, 1–61. Toronto, Canada: Association for Computational Linguistics. <https://aclanthology.org/2023.iwslt-1.1.pdf>.
- Armstrong, Stephen, Andy Way, Colm Caffrey, Marian Flanagan, Dorothy Kenny, and Minako O'Hagan. 2006. “Improving the quality of automated DVD subtitles via example-based machine translation.” In *Proceedings of Translating and the Computer* 28, 1–13. London: Aslib. <https://aclanthology.org/2006.tc-1.9.pdf>.
- Bywood, Lindsay, Panayota Georgakopoulou, and Thierry Etchegoyhen. 2017. “Embracing the threat: Machine translation as a solution for subtitling.” *Perspectives* 25 (3): 492–508. <https://doi.org/10.1080/0907676X.2017.1291695>.

- Deryagin, Max, Miroslav Pošta, and Daniel Landes. 2021. *AVTE Machine Translation Manifesto. Audiovisual Translators Europe*. [https://avteurope.eu/wp-content/uploads/2021/09/Machine-Translation-Manifesto\\_ENG.pdf](https://avteurope.eu/wp-content/uploads/2021/09/Machine-Translation-Manifesto_ENG.pdf).
- Díaz Cintas, Jorge. 2005. "The ever-changing world of subtitling: Some major developments." *Research on Translation for Subtitling in Spain and Italy*, edited by John D. Sanderson, 17–26. Universidad de Alicante. [https://spiral.imperial.ac.uk/bitstream/10044/1/1401/1/Alicante\\_05.pdf](https://spiral.imperial.ac.uk/bitstream/10044/1/1401/1/Alicante_05.pdf).
- Díaz Cintas, Jorge, and Aline Remael. 2007. *Audiovisual Translation: Subtitling*. Manchester: St. Jerome.
- Díaz-Pérez, Francisco Javier. 2020. "Translating swear words from English into Galician in film subtitles. A corpus-based study." *Babel* 66 (3): 393–419. <https://doi.org/10.1075/babel.00162.dia>.
- Etchegoyhen, Thierry, Lindsay Bywood, Mark Fishel, Panayota Georgakopoulou, Jie Jiang, Gerard van Loenhout, Arantza del Pozo, et al. 2014. "Machine translation for subtitling: A large-scale evaluation." In *Proceedings of the Ninth International Conference on Language Resources and Evaluation LREC'14, May 2014*, edited by Nicoletta Calzolari, Khalid Choukri, Thierry Declerck, Hrafn Loftsson, Bente Maegaard, Joseph Mariani, Asuncion Moreno, et al., 46–53. Reykjavik, Iceland. [http://www.lrec-conf.org/proceedings/lrec2014/pdf/463\\_Paper.pdf](http://www.lrec-conf.org/proceedings/lrec2014/pdf/463_Paper.pdf).
- European Commission. 2008. "Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions: Multilingualism: An Asset for Europe and a Shared Commitment." September 18, 2008. <https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2008:0566:FIN:EN:PDF>.
- . 2010. "Directive 2010/13/EU of the European Parliament and of the Council." March 10, 2010. <https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2010:095:0001:0024:en:PDF>.
- . 2018. "Directive (EU) 2018/1808 of the European Parliament and of the Council." November 14, 2018. <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32018L1808>.
- Europeana. "Europeana subtitled AI pipeline and training suite." Accessed April 20, 2024. <https://pro.europeana.eu/page/europeana-subtitled-ai-pipeline-and-training-suite>.
- Georgakopoulou, Panayota. 2012. "Challenges for the audiovisual industry in the digital age: The ever-changing needs of subtitle production." *The Journal of Specialised Translation* 17: 78–103.
- Gottlieb, Henrik. 1992. "Subtitling – A new university discipline." In *Teaching Translation and Interpreting: Training, Talent and Experience. Papers from the First Language International Conference, Elsinore, Denmark 31 May-2 June 1991*, edited by Cay Dollerup, and Anne Loddegaard, 161–70. Amsterdam: John Benjamins Publishing.
- Kammerevert, Petra. 2023. "REPORT on the Implementation of the Revised Audiovisual Media Services Directive (2022/2038(INI))." European Parliament, Committee on Culture and Education. April 12, 2023. [https://www.europarl.europa.eu/doceo/document/A-9-2023-0139\\_EN.pdf](https://www.europarl.europa.eu/doceo/document/A-9-2023-0139_EN.pdf).
- Karakanta, Alina, Luisa Bentivogli, Mauro Cettolo, Matteo Negri, and Marco Turchi. 2022. "Post-editing in automatic subtitling: A subtitlers' perspective." In *Proceedings of EAMT-Translators' Track, 1–3 June 2022*, edited by Helena Moniz, Lieve Macken, Andrew Rufener, Loïc Barrault, Marta R. Costa-jussà, Christophe Declercq, Maarit Koponen, et al., 261–70. Ghent. <https://aclanthology.org/2022.eamt-1.29.pdf>.
- Koletnik Korošec, Melita. 2011. "Applicability and challenges of using machine translation in translator training." *ELOPE: English Language Overseas Perspectives and Enquiries* 8 (2): 7–18. <https://doi.org/10.4312/elope.8.2.7-18>.
- Mattsson, Jenny. 2006. "Linguistic variation in subtitling: The subtitling of swearwords and discourse markers on public television, commercial television and DVD." In *MuTra 2006 – Audiovisual Translation Scenarios: Conference Proceedings*, edited by Mary Carroll, Heidrun Gerzymisch-Arbogast, and Sandra Nauert, 47–56. Copenhagen. [http://www.euroconferences.info/proceedings/2006\\_Proceedings/2006\\_Mattsson\\_Jenny.pdf](http://www.euroconferences.info/proceedings/2006_Proceedings/2006_Mattsson_Jenny.pdf).
- Media Consulting Group. 2011. "Study on the use of subtitling. The potential of subtitling to encourage foreign language learning and improve the mastery of foreign languages. EACEA/2009/01." European Commission, Directorate-General Education and Culture. June, 2011. [http://publications.europa.eu/resource/cellar/e4d5cbf4-a839-4a8a-81d0-7b19a22cc5ce.0003.01/DOC\\_1](http://publications.europa.eu/resource/cellar/e4d5cbf4-a839-4a8a-81d0-7b19a22cc5ce.0003.01/DOC_1).

- Mezeg, Adriana. 2023. "Ali sploh še potrebujemo prevajalce?" *Ars et humanitas* 17 (1): 139–54. <https://doi.org/10.4312/ars.17.1.139-154>.
- Mohar, Tjaša, Sara Orthaber, and Tomaž Onič. 2020. "Machine translated Atwood: Utopia or dystopia?" *ELOPE: English Language Overseas Perspectives and Enquiries* 17 (1): 125–41. <https://doi.org/10.4312/elope.17.1.125-141>.
- Petukhova, Volha, Rodrigo Agerri, Mark Fishel, Sergio Penkale, Arantza del Pozo, Mirjam Sepesy Maučec, Andy Way, et al. 2012. "SUMAT: Data collection and parallel corpus compilation for machine translation of subtitles." In *Published Scientific Conference Contribution*, 21–28. [http://www.lrec-conf.org/proceedings/lrec2012/pdf/154\\_Paper.pdf](http://www.lrec-conf.org/proceedings/lrec2012/pdf/154_Paper.pdf).
- Poibeau, Thierry. 2022. "On 'Human Parity' and 'Super Human Performance' in machine translation evaluation." In *Proceedings of the Thirteenth Language Resources and Evaluation Conference*, edited by Nicoletta Calzolari, Frédéric Béchet, Philippe Blache, Khalid Choukri, Christopher Cieri, Thierry Declerck, Sara Goggi, et al., 6018–23. Marseille, France: European Language Resources Association. <https://aclanthology.org/2022.lrec-1.647>.
- . 2023. "How machine translation changed the translation industry." *The Responsible AI Forum*, TUM, Sep 2023, Munich (Allemagne), Germany. <https://hal.science/hal-04242026>.
- Popowich, Fred, Paul McFetridge, Davide Turcato, and Janine Toole. 2000. "Machine translation of closed captions." *Machine Translation* 15: 311–41. <https://www.jstor.org/stable/20060451>.
- Reiner, Rob, dir. 2014. *And So It Goes*. Škofljica: Blitz Film & Video Distribution. DVD.
- Sepesy Maučec, Mirjam, Marko Presker, Danilo Zimšek, Matej Rojc, Damjan Vljaj, Darinka Verdonik, and Zdravko Kačič. 2012. "Izdelava slovensko-srbskega vzporednega korpusa podnapisov za razvoj strojnega prevajanja v projektu SUMAT." In *Zbornik Osmе konferenсe Jezikovne tehnologije, 8. do 12. oktobra 2012, [Ljubljana, Slovenia] = Proceedings of the Eighth Language Technologies Conference, October 8th-12th, 2012, Ljubljana, Slovenia: zbornik 15. mednarodne multikonference Informacijska družba - IS 2012, zvezek C; proceedings of the 15th International Multiconference Information Society - IS 2012, vol. C*, 167–72. Ljubljana. <http://nl.ijs.si/isjt12/JezikovneTehnologije2012.pdf>.
- Simič, Nina. 2022. "Googlov angleško slovenski prevajalnik ni več najboljši." *Moje finance*, June 30, 2022. Accessed December 27, 2023. <https://mojefinance.finance.si/9001658/prevajalniki?src=undefined>.
- Smith, Chris A. 2023. "Productivity from a metapragmatic perspective: Measuring the diachronic coverage of the low level lexico-grammatical construction Have the N (Body Part/Attitude) to ↔ Metapragmatic Comment> using the COHA." *Languages* 8 (2): 92: 1–23. <https://doi.org/10.3390/languages8020092>.
- Sonix. "Sonix.ai." Accessed December 27, 2023. <https://sonix.ai>.
- Tiedemann, Jörg. 2007. "Improved sentence alignment for movie subtitles." In *Recent Advances in Natural Language Processing RANLP 2007*, edited by Galia Angelova, Kalina Bontcheva, Ruslan Mitkov, Nicolas Nicolov, and Nikolai Nikolov, 582–88. Borovets: Bulgaria.
- Trupej, Janko. 2015. "Strategije za podnaslovno prevajanje profanosti v slovenščino." *Slavistična revija* 63 (1): 17–28. <https://www.dlib.si/details/URN:NBN:SI:DOC-0XTSKK8O>.
- Volk, Martin. 2009. "The Automatic Translation of Film Subtitles. A Machine Translation Success Story?" *Journal for Language Technology and Computational Linguistics* 24 (3): 115–28.
- Volk, Martin, and Søren Harder. 2007. "Evaluating MT with translations or translators. What is the difference?" In *Proceedings of Machine Translation Summit XI: Papers*, edited by Bente Maegaard, 499–506. Copenhagen, Denmark. <https://aclanthology.org/2007.mtsummit-papers>.

# **Part V**

## **Literature**



Michelle Gadpaille  
University of Maribor, Slovenia

2024, Vol. 21 (1), 211-227(316)  
journals.uni-lj.si/elope  
<https://doi.org/10.4312/elope.21.1.211-227>  
UDC: 821.111.09-31"18":2-557

## The Discourse of the “Good Death” in Nineteenth-Century Fiction and a Historical Case Study

### ABSTRACT

What people said on their deathbeds preoccupied 19th-century Protestants. Fear of the void after death or of punishment in the afterlife made death terrifying. The process of dying, however, was ritualized to negate these terrors. Having a “good death” constituted reassurance for the dying, comfort for the family, and a lesson for the community. Such comfort was available even from fictional deathbed scenes, where sentiment and imagery softened the harsh reality. This interdisciplinary study explores the similarities between real and fictional deathbed narratives, using a sample of 19th-century novels, two historical deathbeds from Scotland, and a case study of an actual deathbed from the West Indies. The pattern of the good death – in both fiction and biography – includes specific elements such as witness, testimony, requests and advice, music, scripture, and confident hope of Heaven. Some deathbeds even became virtual stages or pulpits, with the dying person at the centre.

**Keywords:** Victorian deathbeds, Presbyterian missions, Victorian novels, colonial Jamaica, Eliza Hitchcock Davidson

### Diskurz »dobre smrti« v pripovedni prozi 19. stoletja in zgodovinska študija primera

#### IZVLEČEK

Protestanti so v 19. stoletju veliko pozornosti namenjali besedam, izrečenim na smrtni postelji. Ljudje so se bali smrti zaradi strahu pred praznino ali kaznijo v posmrtnem življenju. Strahove so poskusili pregnati z ritualizacijo postopka umiranja. “Dobra smrt” naj bi pomirila umirajočega, potolažila družino in predstavljala nauk za skupnost. Tovrstno tolažbo so ponujali celo prikazi umiranja v fikciji, kjer so kruto stvarnost omehčala čustva in podobe. Pričujoča interdisciplinarna raziskava ugotavlja podobnosti med resničnimi in fiktivnimi prikazi umiranja v romanih 19. stoletja: na primeru dveh škotskih zgodovinskih romanov ter na primeru resnične, zgodovinsko izpričane smrti v Zahodni Indiji. Vzorec dobre smrti – tako v leposlovju kot v biografiji – vključuje posebne elemente, kot so priče, pričevanje, prošnje in nasveti, glasba, sveto pismo in upanje na pot v nebesa. Nekateri opisi ljudi na smrtni postelji so se spremenili celo v navidezne odre ali prižnice, na katerih je bila v ospredju umirajoča oseba.

**Ključne besede:** umiranje v viktorijanski dobi, prezbiterijanski misijoni, viktorijanski romani, kolonialna Jamajka, Eliza Hitchcock Davidson

# 1 Introduction: The Signifying Death

Deathbeds were once communal scenes of testament, witnessed by family members and surrounded by pastoral care (Masur 2015, 12; Bross 2001, 332). Moreover, what the dying person said in their final hours was once considered to have serious consequences for their soul and was thus worthy of being recorded. For example, Puritan deathbed testaments from Martha's Vineyard were explored by Sarah Rivett, to reveal "a written record of divine translation from death to life, from the visible to the invisible world, from the dark glass of limited perception to unfiltered revelation" (2017, 472). Such translation of an event into doctrinal fulfilment made the articulate death a valuable motivator for the living, while it assured a happy afterlife for the dying. When Lisa Shaver studied memoirs and obituaries from 19<sup>th</sup>-century Methodist deaths, she found that one function of women's "deathbed pulpits" was the motivation and instruction of the living (Shaver 2008, 20; Zaaraoui 2017, 14). According to Kristina Bross, deaths of "Indians" in Puritan North America were always interpreted as a sign from God (Bross 2001, 325), and converts rapidly learned the deathbed speech conventions that would assert their "Christian identity" to a sceptical community (Bross 2001, 332). Professions of faith on one's deathbed could mark one as a "visible saint" (Bross 2001, 332) and thus were valuable currency for both missionaries and their dying converts. Deathbed speeches fulfilled multiple functions (Shaver 2008, 20), and often featured prescribed rhetorical strategies, actions, and gestures.

The poignancy of a speaking death was enhanced when the person was female or a child (Masur 2015, 40-41; Cassidy 2002, 210, 213, Rivett, 2011, 17; Martineau 1844, 104-5). In the case of a child, elaborate professions of faith were not expected, but even minor utterances or gestures took on significance to the survivors. Churchgoing women, however, were expected to make what was called a "good death", in the process achieving the social dominance (though fleeting) denied them in life.

To 21st-century sensibilities, the fascination with the process of dying can seem morbid, even when its prominence arises from its indexical function as pointing to the future of the dying person's soul.<sup>1</sup> Nineteenth-century readers loved a saintly invalid and relished reading of an eloquent demise, whether in a memoir of a neighbour or fellow congregant (White 1985, 5; Bross 2001, 332), or in the case of literary heroines (Zaaraoui 2017, 13; White 1985, 5), where the aesthetics of death prevailed. "Do not the most common-place writers of fiction crowd their novels with death scenes . . .? and do not those who stay home learn all they can of the last words and demeanour of the sufferers?" asked Harriet Martineau from her sickroom in 1844 (1844, 104). Fictional deathbeds allowed for a form of "shared, public mourning" (White 1985, 5) among the reading community. Among the earliest deathbed narratives in the English novel is that of the heroine in Samuel Richardson's *Clarissa* (1748), which falls outside the scope of this article.<sup>2</sup> Prime examples of popular fictional deaths include those

---

<sup>1</sup> Interest in literary deathbed scenes has not, however, completely disappeared in the 21st century. See *The Guardian's* 2009 feature "Ten of the Best Deathbed Scenes in Literature". In a contemporary update to the issue, Heather Duncan raised the question of potential "death practices" (2018, 95) for the 21st-century, once digitalization makes a personal afterlife a serious possibility, not a matter of faith, as with the spiritual afterlife sought on Victorian deathbeds.

<sup>2</sup> Clarissa's deathbed resembles the 19th-century pattern in being emotional, pious, witnessed, and self-directed. Having accepted her fate, Clarissa makes an extended exit, sending copious letters (Richardson 1863, 297), requesting pillows

of Helen Burns in *Jane Eyre* (1847), Beth March from *Little Women* (1868), Little Eva from Harriet Beecher Stowe's *Uncle Tom's Cabin* (1852), and, best known of all, Little Nell from Charles Dickens' *The Old Curiosity Shop* (1840–41).

I will examine some of these literary deathbed scenes for their narrative tropes and rhetorical strategies. These fictional deathbeds will be juxtaposed with two historical deathbeds recorded by biographers, and finally with one extended deathbed scene from the mid-19<sup>th</sup> century, to explore the differences in the aestheticization of death, ritualization, and the notion of the deathbed as a staged spectacle (Zaaraoui 2017, 13).

## 2 Fictional Deathbeds

The deathbed scene in fiction did not occupy the cultural foreground until near the middle of the 19th century. Fictional characters did die in earlier novels, but with less emphasis upon the deathbed as a separate tableau with a motivational function. Deathbed scenes from Mary Shelley's novel *The Last Man* (1826) are instructive in this regard. This work imagines a plague pandemic, originating in the east and spreading to England. Despite the innumerable deaths in this novel, few are foregrounded as individual losses. Even when treated in detail, these early deathbeds lack the evangelical tone of fictional deathbeds a generation later.

Writing in the 1820s and out of the Romantic movement, Shelley offers nothing like the sentimental deathbed of the mid-Victorian age. *The Last Man* belongs to an indeterminate genre, part science fiction, part Gothic melodrama, in which proselytizing Christian belief plays no part. The novel's many deaths generally occur off-stage or are briefly sketched. Here, for instance is the death of the Countess of Windsor, who is among the last survivors: "In the morning we had seen her apparently in health – in the evening, Lucy, before we retired to rest, visited our quarters to say that she was dead" (Shelley 1994, 414).

More extensive description marks the death of Evadne earlier in the novel, when the plague has not yet taken hold. Hers constitutes the antithesis to the "good death" that will be illustrated later. First, it is not a deathbed at all, since it takes place outdoors on an abandoned battleground, a piece of "corse-strewn earth" (Shelley 1994, 180) and is thus beyond the realm of the domestic. The good death optimally occurred at home, within the domestic sphere (Zaaraoui 2017, 13). Second, the dying person, though dressed as a soldier, is a woman, a Greek princess, and the doomed lover of Raymond, one of the novel's protagonists. Next, there is the attitude of the dying woman: far from offering blessings to witnesses, Evadne calls down vengeance on Raymond for having deserted her:

Many living deaths have I borne for thee, O Raymond, and now I expire, thy victim! – By my death I purchase thee – lo! The instruments of war, fire, the plague are my servitors . . . Fire, and war, and plague, unite for thy destruction – O my Raymond, there is no safety for thee! (Shelley 1994, 181).

---

(297), offering blessings and forgiveness (302–4), and issuing orders about the disposal of her body (298). This last constitutes the real distinction between *Clarissa* and its successors: Richardson's dying heroine is flesh and blood, described as a "corpse" and needing a coffin, a hearse, and aromatic herbs to mask the odour (1863, 326–27). Such physical frankness is absent from the 19<sup>th</sup>-century examples examined in section 2 "Fictional Deathbeds".

She briefly anticipates meeting Raymond in Heaven (Shelley 1994, 181), but otherwise her dying address comprises equal parts prophecy and threat. Hers is not a soul-comforting death, nor an instructive moral lesson. Instead, this is a gothic death, filled with disguise, wilful passion, transgressive roles, and an intensity of earthly emotion that will be absent from the Victorian deathbeds explored in the next section. With her Romantic sensibility, Shelley emphasizes individual self-determination over any surrender to the arms of a Christian Jesus.

The Romantic aspect of Shelley's death scenes becomes even more apparent in the scene where little Evelyn dies, near the end of the novel. The death of a child or adolescent would later become the ideal nucleus for the "good death", but Shelley's child, though cherished, dies in a natural and not a moral universe. The son of Lionel, the narrator, Evelyn is among the band of English plague survivors that has fled to Italy (Shelley 1994, 434). Given his youth, he could have represented the last hope for humanity's continuation. Unfortunately, he is "seized with sudden fever" (Shelley 1994, 434) and dies of typhus within a fortnight. His is a tender deathbed, as carefully witnessed as any of the mid-Victorian ones, but *what* is witnessed differs substantially:

His little form and tiny lineaments encaged the embryo of the world-spanning mind of man. Man's nature, brimful of passions and affections, would have had an [sic] home in that little heart, whose swift pulsations hurried towards their close. His small hand's fine mechanism, now flaccid and unbent, would in the growth of sinew and muscle, have achieved works of beauty or of strength. His tender rosy feet would have trod in firm manhood the bowers and glades of earth. (Shelley 1994, 434–35)

On display here are Shelley's own feelings of maternal loss, her interest in natural philosophy and in the capacity of human beings to feel passion, appreciate natural beauty and add creatively to that beauty. The father's deathbed wishes do not concern an afterlife but comprise regret that Evelyn has been denied the chance of a fulfilling life.

This child's passing conveys no message of salvation, an absence that the narrator notes explicitly: "I have heard that the sight of the dead has confirmed materialists in their belief. I ever felt otherwise. Was that my child – that moveless decaying inanimation?" (Shelley 1994, 435). In Shelley's novel, even the death of a child does not become a didactic spectacle enacting the drama of faith fulfilled.

Over the next generation, however, the deathbed agenda changes between Shelley's Romantic era and the Victorian atmosphere of domestic piety in which Charles Dickens serialized his novels. I now turn to deathbed scenes from major novels of the 1840s and later, starting with three that epitomize the pattern of the good death in literature.

Since the death scene of Little Nell in Dickens' *The Old Curiosity Shop* has sometimes been taken as the pattern of the Victorian deathbed (Masur 2015, 75), I begin with an analysis of its components.

Like other good deathbeds, Nell's is witnessed; there are people "all about her at the time" because her death is expected (Dickens 1868, 324–27). Second, it takes place at a temporally

liminal time – soon after daybreak. The dying individual both receives messages and utters blessings. Nell is too young to preach on her deathbed, so a simple “God bless you” suffices. The good deathbed is also a lucid one. Gestures include the request for an embrace or kiss, a smile, and the intimation that such a facial expression is unique and angelic. Dickens gives Little Nell a restrained death scene, with little melodrama and no gore, although much sentiment. The moment of death has passed when the reader is admitted to the death chamber, and much of what we think we know about Little Nell’s death is due to George Cattermole’s woodcut illustration of the scene.<sup>3</sup> Cattermole shows Nell lying under the coverlet in a bed with an elaborate headboard. With no physical indications of disease or suffering, she looks tranquil and older than her years. Her eyes are closed, and her right hand clasps a book. The room has arched windows and faint indications of drapery. The most striking element in the illustration is the carved headboard, which rises over Nell’s head as if it were a dream emanating from her. It depicts the Virgin Mary with the infant Christ on her lap; it is thus the illustrator and not the writer who introduces the concept of sacrificial death.

Another best seller from the mid-19th century was Harriet Beecher Stowe’s anti-slavery novel, *Uncle Tom’s Cabin* (1852). This popular account of life on a Louisiana plantation created a cluster of stereotypes: Topsy, Little Eva and Uncle Tom himself. The atmosphere of the novel drips with Christian fervour, and even so young a child as Eva St Clare knows that she will die soon: “‘Uncle Tom,’ said Eva, ‘I’m going there’ . . . ‘*I’m going, before long*’” (Stowe 1994, 227). Although this remark comes in Chapter XXII, the actual deathbed occurs four chapters later, in XXVI, titled “Death”. The scene is Eva’s spacious bedroom (Stowe 1994, 246), to the décor of which Stowe devotes a lengthy passage. The scene is reminiscent of Cattermole’s illustration of Little Nell’s death. Stowe clearly constructs Eva’s room as a shrine: there is a “statuette of Jesus”, and a “sculptured angel with drooping wings” along with rose-buds, rose-coloured damask and a great deal of marble and alabaster (Stowe 1994, 247). Although Uncle Tom brings a floral offering to the Jesus statuette each morning, this shrine is the home to a more recent saint in the form of the dying child. It is for Eva that the drooping angel holds out “a crown of myrtle leaves” (Stowe 1994, 247). Eva is fully conscious of her plight and takes charge of the deathbed, giving orders freely: “I want to see all our people together. I have something I *must* say to them” (Stowe 1994, 250). Here, “our people” means the enslaved population of the estate, and Uncle Tom is present as a prime witness to this most public of fictional deaths. In a most un-childlike way, Eva preaches from her deathbed:

If you love me, you must not interrupt me so. Listen to what I say. I want to speak to you about your souls . . . Many of you, I am afraid, are very careless. You are thinking only about this world. I want you to remember that there is a beautiful world, where Jesus is. (Stowe 1994, 251)

In the middle of this outpouring, Stowe gives the child a sudden realization that reading the Bible is unavailable to these enslaved people as a means to salvation: “O, dear! You can’t read – poor souls” and she hid her face in the pillow and sobbed . . .” (Stowe 1994, 251). Stowe

<sup>3</sup> The illustration was titled “At Rest (Nell dead)” and appeared in *The Old Curiosity Shop*, part 39 on 30 January 1841 when it was serialized in the journal *Master Humphrey’s Clock*. A reproduction is available at <https://victorianweb.org/art/illustration/cattermole/9.html>

foregrounds her main anti-slavery message. Through the scene, the dying child advances to a position of almost adult conversational privilege (Stowe 1994, 251–55) before the release comes. The change is initially visible on Eva's face where "there was no ghastly imprint, – only a high and almost sublime expression" (Stowe 1994, 256). By the next page, her final words appear:

The child lay panting on her pillows, as one exhausted, – the large clear eyes rolled up and fixed. Ah, what said those eyes, that spoke so much of heaven? Earth was past, and earthly pain; but so solemn, so mysterious, was the triumphant brightness of that face, that it checked even the sobs of sorrow. They pressed round her, in breathless stillness.

"Eva," said St. Clare, gently.

She did not hear.

"O, Eva, tell us what you see! What is it?" said her father.

A bright, a glorious smile passed over her face, and she said, brokenly, – "O! love, – joy, – peace!" gave one sigh, and passed from death unto life! (Stowe 1994, 257)

Stowe's Eva, therefore, though a powerless child, is made "powerful by selflessly dying for others" (White 1985, 7, 9). This release of social power through the spectacle of the "good death" will prove to be traceable in the historical case study of Eliza Hitchcock Davidson's death.

Almost a generation later, Louisa May Alcott described the death of Beth March, one of the sisters in *Little Women* (1868). Alcott titled the chapter "The Valley of the Shadow" (1915, 522), referencing Psalm 23. As with Nell and Eva, the death is anticipated, the family has "accepted the inevitable" (Alcott 1915, 520) and set up a room for Beth's last days. Like Nell and Eva, though older, Beth is given saintly attributes, for she is like "a household saint in its shrine" (Alcott 1915, 521). Though her body is described as "feeble" and a "wreck" (Alcott 1915, 522), Beth has no grisly physical symptoms. Instead, Alcott highlights the contrast between body and soul, the latter being both serene and "strong" (1915, 522). Though cast as a saint, Beth is additionally imaged as a pilgrim ("the first pilgrim called was likewise the fittest" (Alcott 1915, 522)), in obvious reference to one of the sisters' favourite books, *The Pilgrim's Progress*, which lies upon the table next to Beth's deathbed (Alcott 1915, 524).

More articulate than Nell and less bossy than Eva, Beth declares that she has learned to face death without fear, and even with happiness: "I don't fear it any longer", she avers to her sister Jo (Alcott 1915, 526). One unusual feature of this deathbed is the poem written by Jo (Alcott 1915, 524–25). Found within the pages of *The Pilgrim's Progress*, the poem points to Beth's "serene and saintly presence", while confirming that Jo feels the painful experience has taught her a valuable life lesson: "And while learning this hard lesson, / My great loss becomes my gain. / For the touch of grief will render / My wild nature more serene." Thus, Beth's deathbed acquires agency among the living. Moreover, having read the poem Beth is reassured about her quiet, domestic life: "Then I don't feel as if I'd wasted my life. I'm not so good as you make me, but I have tried to do right" (Alcott 1915, 525).

Alcott's deathbed scene overtly alludes to the literary trope of the articulate deathbed: "Seldom except in books do the dying utter memorable words, see visions, or depart with beatified countenances, and those who have sped many parting souls know that to most the end comes as naturally and simply as sleep" (Alcott 1915, 526). Beth's end comes "in the dark hour before the dawn" (Alcott 1915, 526) and is compared to the going out of the tide. There are no parting words but "one loving look, one little sigh" (Alcott 1915, 527).

Despite having set aside the literary deathbed in a lightly mocking tone, Alcott cannot completely leave it behind; Beth's deathbed scene conforms to much of the pattern: the liminal departure time, the lucid messages, the didactic effect of the death on surviving family members, and the deployment, not here of music, but of poetry and didactic literature.

The deathbed scene in Ellen Woods's sensation novel *East Lynne* (1861) provides a contrast. A small boy, William Carlyle, has been ill for a while, and the doctor has even ordered his lessons to be stopped because "he'll never want them" (Wood 2000, 582). The disease affects his lungs (2000, 623) and is probably consumption (i.e. tuberculosis): "The brilliant hectic flush had gone from his cheeks, his features were white and wasted, his eyes large and bright" (Wood 2000, 641). This deathbed scene is notable for its length and the fact that the narrator's attention is not solely on the dying child, but equally on his main attendant, who is his disguised mother, Lady Isabel. The second notable feature is the medicalization of this death; in contrast to Nell in Dickens, bodily symptoms are described in detail, including ague, sweating and paroxysms (Wood 2000, 643–44). There is emphasis on the "hollow breath" and the "blue, pinched, ghastly look" (Wood 2000, 649). William Carlyle emerges as a suffering body. Despite the materiality of this deathbed, some emphasis is placed on the dying boy's destination after death. He is assured repeatedly that he will be going to Heaven:

There will be the beautiful city with its gates of pearl, and its shining precious stones, and its streets of gold; and there will be the clear river, and the trees with their fruits and their healing leaves, and the lovely flowers; and there will be the harps, and music, and singing. (Wood 2000, 641)

As with evangelical deathbeds, this child is confident in the coming embrace of Jesus:

"Madame Vine, will Jesus come for me, do you think, or will he send an angel?"

"Jesus has *promised* to come for his own redeemed – for those who love Him and wait for him." (Wood 2000, 641–42)

Nevertheless, he is not uniformly joyous, for he is anxious to establish whether his biological mother will be there in Heaven, too: "[D]o you think mamma will be there?" (Wood 2000, 642).

Overall, then, a happy afterlife, united with his family, takes its importance not from doctrinal or spiritual certainty, but from the exigencies of the plot. William is made to yearn for family unification in Heaven, so that the disguise worn by Lady Isabel will slip and thus advance the plot. Eventually, William dies while "Madame Vine" (his mother) is paying no attention but is wrapped in her own trauma, "Lost in thought, in anguish past and present, in self-condemning repentance" (Wood 2000, 652). It falls to the nurse to discover that the boy has finally died:

Joyce, advancing with a quiet step drew aside the clothes to look at William. “Master says he has been wanting me,” she observed. “Why – oh!”

It was a sharp, momentary cry, subdued as soon as uttered. Madame Vine sprang forward to Joyce’s side, looking also. The pale young face lay calm in its utter stillness; the busy little heart had ceased to beat. (Wood 2000, 652)

After the boy’s death, attention shifts to the indulgent grief of his mother. Unlike Little Nell in Dickens’ novel, the child’s death has no function in *East Lynne* other than as a catalyst for the revelation of Lady Isabel’s return, disguised as the family governess. Her emotional outpouring might once have constituted atonement for family abandonment. However, her remorse for her actions far outweighs her grief at the boy’s loss. Unlike Nell, Beth and Eva, William is neither saintly, ethereal nor angelic, and his deathbed offers no doctrinal reassurance for the reader: the boy does not profess his faith, nor ask for hymns or lines from scripture. And although there is always at least one person with the child, his death is in no sense witnessed; not even his mother is paying attention in his final moment.

### 3 Historical Deathbed Narratives

Before turning to the historical case study, I will establish the context by looking at two actual death scenes that were recorded by a third person. Both took place in the mid-19<sup>th</sup> century and in Scotland. Unlike the Jamaican case study, both involve dying *men*, and men of consequence within their church and community.

The earlier of the two biographies narrating a deathbed scene is that of Robert Murray M’Cheyne (1813–1843), a minister in the church of Scotland, who died in 1843 at the age of 30, never having been in robust health (Millar 1900). He worked as a missionary and was famed as an ardent preacher (Millar 1900). The account of his deathbed was written by his friend Andrew Bonar, who had accompanied M’Cheyne on a mission to Palestine in 1839. Bonar narrates how his friend insisted on conducting regular pastoral duties until he “felt chilled and unwell” (1848, 144). “He believed that he had taken the fever, and it was so. That night he lay down upon the bed from which he was never to rise” (Bonar 1848, 145). M’Cheyne accepts his approaching death, while quoting from Psalm 126 (undoubtedly thinking of its closing lines: “They that sow in tears shall reap in joy. / He that goeth forth and weepeth, bearing precious seed, shall doubtless come again with rejoicing, bringing his sheaves with him” (Bonar 1848, 145). To this, M’Cheyne adds verses from Matthew 11:28: “Come to me, all who labor and are heavy laden, and I will give you rest” (Matthew 11: 28–30). There are the expected joyful exclamations from this deathbed, in which we glimpse the oratory for which M’Cheyne was famed:

When his servant entered the room again, he exclaimed with a joyful voice, “My soul is escaped as a bird out of the snare of the fowler; the snare is broken, and I am escaped.” His countenance, as he said this, bespoke inward peace. Ever after he was observed to be happy. (Bonar 1848, 145)

Since the young preacher had a “refined musical taste” (Millar 1900), it is unsurprising that hymns featured in his dying moments:

On Tuesday (the 21st) his sister repeated to him several hymns. The last words he heard, and the last he seemed to understand, were those of Cowper's hymn, *Sometimes the light surprises the Christian as he sings*. (Bonar 1848, 146)<sup>4</sup>

M'Cheyne succumbs to delirium at the end, but even in this he quotes scripture, 1 Corinthians 15:58, "dwelling with much emphasis on the last clause, *"forasmuch as ye know that your labor is not in vain in the Lord"* (Bonar 1848, 146). Otherwise, his delirium takes the form of loud prayers for his congregation. His biographer stresses the "happy frame" (Bonar 1848, 146) of the dying man along with his care for those he was leaving behind.

M'Cheyne dies in the morning, after having made a symbolic gesture: he "lifted up his hands as if in the attitude of pronouncing the blessing, and then sank down". There is no indication of suffering at the moment of death: "Not a groan or a sigh, but only a quiver of the lip" (Bonar 1848, 146).

M'Cheyne's deathbed resembles fictional ones in its uniformly positive tone. His biographer acknowledges delirium, so the dying man is not always lucid; nevertheless, even in delirium M'Cheyne spoke as a preacher. He never lost his sense of his role in his community. The deathbed is closely attended, lovingly witnessed, and recorded in detail, down to the chapter and verse of the hymns and gospels that he recited. Additionally, his discourse includes a lesson for the living ("You must be awakened in time, or you will be awakened in everlasting torment, to your eternal confusion" (Bonar 1848, 146)). Despite M'Cheyne's youth, he is not ascribed any angelic appearance or divine aura. He dies as what he was – a committed member of the Church of Scotland, secure in his own salvation.

A later deathbed scene, that of William Cunningham (1805–1861), was recorded in detail by Robert Rainy and James Mackenzie in 1871 (*Life of William Cunningham, D.D Principal and Professor of Theology and Church History, New College, Edinburgh*). This 1861 death deserved notice because Cunningham had been a famous theologian (Rainy and Mackenzie 1871, 222 ff.) and founder of the Free Church of Scotland (240 ff.). This deathbed scene conforms well to the fictional pattern. The main feature of this deathbed is its sense of active agency: Cunningham sends messages, makes a will and bequests, delegates functions, and counsels his visitors:

He gave particular directions as to all necessary things he wanted done, even to the inscription on his tomb-stone, and the persons he wished to be present at his funeral; and said much to soothe and comfort those he was leaving behind him. (Rainy and Mackenzie 1871, 477)

In contrast to some fictional deathbeds, Cunningham's is one where interaction with the world is far from over. As with M'Cheyne, Cunningham finds comfort in specific doctrines; he asked for his Bible, and two books: *The Confession of Faith*<sup>5</sup> and John Newton's the *Olney*

<sup>4</sup> The hymn's actual opening lines are as follows: "Sometimes a light surprises / The Christian while he sings; / It is the Lord who rises / with healing in his wings" (Newton 1825, 256–57). The hymn's title is "Joy and Peace in Believing", and it is numbered XLVIII in this edition.

<sup>5</sup> *The Westminster Confession of Faith* was standard for Church of Scotland and Presbyterian adherents.

*Hymns* (Rainy and Mackenzie 1871, 475). Because of his weakness, the hymns were read to him “to which he listened with great pleasure” (Rainy and Mackenzie 1871, 476). Although he had hoped to sing them, this is prevented by his wife, who thinks the exertion unwise (Rainy and Mackenzie 1871, 475). The *Olney Hymns* are an oddly simple selection for a theologian, full of “Evangelical commonplaces” (Hartley 1949, 221), representing the lowest common denominator of expressed faith. Additionally, Cunningham asserts confidence in his destination: “We shall meet at the right hand.” The biographer admits to some slight “wandering” in Cunningham’s mind towards the end but maintains that “even then at times he spoke quite collectedly” (Rainy and Mackenzie 1871, 477). His “last articulate words” were “I am going home quietly” (Rainy and Mackenzie 1871, 477).

#### 4 Religion in Post-Emancipation Jamaica

As I turn to the main historical deathbed, that of Eliza Hitchcock Davidson, some socio-historical background becomes necessary because Eliza died in the British colony of Jamaica in 1859. At the time of her death, the sugar colony had put slavery in its past (Abolition Act 1834), but already had a population dominated by people of African descent and of mixed race.<sup>6</sup> Eliza belonged to the thin stratum of urban middle-class whites and to the Free Church of Scotland, one of multiple denominations on the island.

One missionary, the Reverend James Phillippo, left an account of the spiritual state of the island’s people, among whom the initial upsurge in church membership after Emancipation had given way to apathy and even backsliding in some rural congregations. Baptist mission churches like Phillippo’s struggled to keep their congregants in line and their home churches supportive of the mission. What Phillippo witnessed among the post-Emancipation population where death was concerned provides context for the account of Eliza Hitchcock Davidson’s death.

At the end of a long career devoted to the freed people of Jamaica, Phillippo testifies to the particular devotion of women to the church (1843, 147). Nevertheless, he reports being distressed by “revolting” funeral practices among the freed slaves (Phillippo 1843, 94); these un-Christian practices start with drumming, extend to dancing and drunkenness, and end with the practice of corpse prophecy (Phillippo 1843, 94). In other words, Phillippo registers the ways in which religion in rural Jamaica was becoming syncretic, moving away from the practices of the home church. Death and dying are singled out by Phillippo as significant elements for assessing the maintenance of doctrinal orthodoxy among the population, because death is the time “when the reality of religion is brought to the test, and no where is it more severely tested than in a land where sickness so often terminates fatally and with so little warning” (Phillippo 1843, 147).

Phillippo devotes considerable attention to accounts of “negro” deathbeds (1843, 149–50). He records that dying “negroes” exhibit “a tranquillity which death could not ruffle,

---

<sup>6</sup> See Chapter 1 (1–22) of Michelle Gadpaille’s *The Ethical Atlantic: Advocacy Networking and the Slavery Narrative, 1830–1850* for a fuller discussion of the issues surrounding abolition in the British colonies.

and a confidence which the king of terrors could not shake" (149). The truly pious in his congregation die with confidence in "their happy prospects" (Phillippo 1843, 159), and he uses these instances as testimony aimed at the living. In the case of one aged female congregant, Phillippo paints a vivid picture of the death chamber: "numbers of persons of all classes successfully crowded around her bed", such that "her chamber seemed the verge of Heaven" (1843, 151). Concomitant with the two Scottish deaths, this one was accompanied by reading of scripture, singing, prayer, and holy conversation (Phillippo 1843, 151). Phillippo mentions "one little black boy" who died imploring his schoolfellows to repent (1843, 151) and reciting poetry, including "Pope's Ode".<sup>7</sup> His account of the death of a female schoolteacher comes closest to that of our case study, Eliza Hitchcock Davidson, for this teacher's death is happy and triumphant, and includes a recitation of the *Nunc dimittis* and Bible readings. Moreover, at least seven later conversions among young people are attributed to this deathbed scene (Phillippo 1843, 153).

Phillippo is thus a biased but credible reporter on the religious situation in Jamaica. His faith and missionary zeal do not cloud his observation of what the congregation really believes and how they behave. The most relevant detail from his account is that women bear the major burden of faith and church service, and the role of women in this context becomes relevant when we turn to Eliza Hitchcock Davidson, who belonged to a branch of the Free Church of Scotland in Jamaica.

## 5 Case Study of a Deathbed: Eliza Hitchcock Davidson<sup>8</sup>

In 1859, a young woman lay dying in a hot room in Kingston, a port in the British colony of Jamaica. Born Eliza Hitchcock, she had married Joseph Davidson, an emigrant from Scotland (Aberlemno, via Glasgow) who, though a businessman, functioned as an elder in the local Presbyterian church. We do not know the cause of Eliza's death, but she was 33 years old in 1859 and had five children.<sup>9</sup> There is, however, no mention of this being a tragedy in childbirth, and she had been in "the very picture of health" (Watson 1859, 150) until a few days prior to her death. She might have fallen victim to a tropical disease, such as yellow fever, an outbreak of which had recently swept through the troops at Newcastle, a military camp in the Blue Mountains above Kingston (Lawson 1859). However, this outbreak of yellow fever occurred too early (in late 1856) to have been the culprit (Lawson 1859). Another potentially deadly foe was cholera, which had visited the West Indies, including Jamaica, between 1850 and 1856 (Jenson and Szabo 2011). Eliza might also have died from malaria, dengue fever, typhoid fever, scarlatina or even from the unknown scourge mentioned by Reverend James Watson at the time: "a fever of an entirely new type for the West Indies" resembling a "malignant typhoid" (Watson 1859, 149).

<sup>7</sup> It is uncertain which of Alexander Pope's poems is meant; none of his works with "ode" in the title seems appropriate to a deathbed. It is likely that the poem meant is Pope's "The Dying Christian to His Soul" which ends with the line "O death where is thy sting?"

<sup>8</sup> Full disclosure: Eliza Hitchcock Davidson is among the author's ancestors.

<sup>9</sup> "She has left behind her five little children, over whom she watched, while living, with the most affectionate solicitude" (Watson 1859, 150).

## 5.1 Revd James Watson as Deathbed Recorder

Although the Davidson family was neither rich nor prominent in the colonial society of Kingston, we know a great deal about Eliza's deathbed because it was recorded by the Revd James Watson and published in the *United Presbyterian Missionary Record* (August 1, 1859). The unusual level of detail in this account of a female deathbed permits us to compare it to the fictional pattern observed in Victorian novels. It also raises the question of why Eliza Davidson's passing deserved such public record.

The elaborate account kept by the Revd Watson over several days shows that Eliza was determined to have the "good death" prized by her co-religionists, called by Watson the "happy death" (1859, 149). Unlike the pious Black Jamaicans whose deathbeds were praised by Revd Phillippo, Eliza had been highly educated in Scotland (Watson 1859, 150) and attended church in Edinburgh under the ministry of Dr Tweedie (150). Eliza's spiritual struggles as a student under Dr Tweedie were recorded by her and made available to the Revd Watson when he was writing up her deathbed:

From the notes of a manuscript book that lies before me, I learn that she had been, while attending a boarding school in Edinburgh, admitted to Dr Tweedie's candidates' class in 1845–46 ... In due time she was admitted a member of the church; and her journal on that occasion, and for some months after, breathes the most ardent and devoted love to the Saviour. (Watson 1859, 150)

Eliza, then, was privately a writer of sorts and took her spiritual notebooks with her even across the Atlantic, although these do not survive.

Writing in *The United Presbyterian Magazine*, the Revd James Watson spoke candidly of the reliance of missionary work on donations from home congregations, and therefore on interesting accounts in missionary magazines to drum up support. As Cheryl Cassidy points out, the obituaries of heathen converts dominated accounts sent back to mission headquarters (2002, 207). "In missions", writes Watson wearily, "as in other things, novelty wears off" (1856, 276). Where once the mission field in a colony like Jamaica had been open for conversion, history had moved on. And what would replace the early accounts of multiple conversions, growing congregations, eradication of heathen practices and the building of new churches on the green hills of Jamaica? Among Baptists, Methodists, and Moravians – all pioneers in the mission field on the island – the situation by the late 1850s had reached a plateau. Emancipation had been achieved, free villages had been founded, churches built, and congregants attracted. But as Phillippo also attested, not every Sunday attendee was a solid conversion. There was backsliding, congregations varied in size, and regular churchgoers tended to be female. Where to find exciting material to fill the pages of the *Missionary Magazine*? Perhaps one answer lay in the riveting deathbed scene of one who, though not a convert, was a literate, and outspoken Presbyterian: enter Eliza Hitchcock Davidson.

## 5.2 Elements of the "good death"

Eliza's deathbed has two layers: first, her deliberate choice of discourse and behaviour, and second, the Revd Watson's shaping of these into a narrative of triumph over death. Together,

they construct the desirable death through the five elements that would later characterize Victorian fictional deathbeds.

### 5.2.1 Witness

As with almost all the deathbeds – fictional or historical – Eliza’s was witnessed: her husband, mother, children and neighbours were there, and the recorder of the scene, the Revd James Watson, was by her bedside intermittently over several days. His act of witness proved particularly important because of his senior role in the Free Church of Scotland and presumed capacity to judge the quality of a “good death”, but also because his account is layered over a reading of Eliza’s own manuscript writings about her faith. The dying Eliza called for everyone to watch and “would say to those who stood by her bed, ‘Come, see how a Christian can die’” (Watson 1859, 151). Watson was sufficiently impressed by this exhortation that he repeated it on the next page. An important set of witnesses were her five children: “The scene of taking farewell of her children was a touching one. She gave them very suitable parting words – spoke and shook hands with each, commending them to God, and telling them not to cry, that she would meet them all again in a better and brighter land” (Watson 1859, 151).

### 5.2.2 Scripture

In parallel with the two historical deaths in Scotland, Eliza used texts from scripture to explain her faith and add comfort to her deathbed: “On the Friday preceding her death she spent almost the whole day in repeating texts of Scripture, interspersed with hymns and paraphrases, and portions of the psalms” (Watson 1859, 151). Eliza’s selections from scripture were diverse, including the Psalms, Isaiah, 2 Timothy, 1 Corinthians (twice) and John (four times). Her deathbed allusions extended to Bunyan’s *The Pilgrim’s Progress* (a choice shared with the sisters from *Little Women*). The amount of scripture remembered, recited, and recorded is remarkable. As a result of this scriptural erudition, the Revd Watson concluded that her death amply demonstrated “the supporting power of the Gospel in view of meeting the last enemy” (1859, 150).

### 5.2.3 Music

Eliza’s musical preferences are fully on show: like her compatriots in the Free Church of Scotland, M’Cheyne and Cunningham, Eliza desired to die to the accompaniment of the sacred music she had loved in life: “Her mind was evidently well stored with some of the most beautiful hymns in our collection. The number, the selectness, the suitableness, of some of these hymns which she repeated were very remarkable and striking... The hymn beginning “I know that my Redeemer lives; what joy the sweet assurance gives!” she repeated to the end” (Watson 1859, 151–52).

Among the hymns that Eliza performed on her deathbed were staples of the hymnal:

“Jesus, I love thy charming name” (Watson 1859, 152)<sup>10</sup>

“The hour of my departure’s come” (Watson 1859, 152)<sup>11</sup>

<sup>10</sup> Philip Doddridge, Hymn 325, in *Hymns Founded on Various Texts in the Holy Scriptures*, 1766.

<sup>11</sup> James Smith, “The Dying Christian”, in *A Selection of Spiritual Poetry*, 1855.

Her preferred hymns came from both the adult and the Sabbath-school hymnals (Watson 1859, 152). The little Sunday-school hymn book was even brought directly into the death chamber by her husband (Watson 1859, 152). The last hymn she recited was one called “Heaven”, taken from a monthly magazine, *The Mother’s Friend* (1848). Like William Cunningham, Eliza wanted to sing but was prevented because of her deteriorating state: “Her friends said she was not able to sing, and it would fatigue and excite her too much. She said, Oh, no. And immediately began, in a sweet, touching, plaintive strain, to sing herself: the rest joined her; and they sung, amid tears, and weeping, and joy, the whole hymn to the end” (Watson 1859, 152). Unlike Cunningham, Eliza had her own way.

#### 5.2.4 Requests and Advice

This deathbed was also marked by a quantity of demands. Eliza’s words have a ring of determination that is difficult to reconcile with her feeble bodily state. “On my first visit to Mrs Davidson, she said very emphatically, ‘Mr Watson, *I am dying*’” (Watson 1859, 150). Eliza took control of the conversation and had to be reminded not to boast about her certainty of salvation: “Speaking of death having no terrors for her now, she was reminded not to boast. Oh no! no! said she, there must be no boasting” (Watson 1859, 151). Having something to prove, she sends for a neighbour to witness the quality of her death:

Then remembering that one of her neighbours had asked her, some weeks before, if she was ready to die, when she replied she hoped so, but she was not sure – remembering that conversation, she had that neighbour immediately sent for. And when she came into the room, she fastened upon her a look of intense earnestness, and said, “Puley, you asked me some months ago if I were ready to die; I was not then, but I am now ... I have sent for you to urge on you special preparation, for it is hard work to die. Remember, Puley, it is not long prayers, it is not going to church, – no, no, it is coming to Jesus.” (Watson 1859, 152)

Looking beyond her neighbour, Eliza also “sent a message to the church and another to the Sabbath school” (Watson 1859, 152), while also commending the pastor’s labours.

#### 5.2.5 “The Triumph of Faith in Christ” (Watson 1859, 149)

Eliza is confident not only of her capacity to face death with dignity and joy, but also of her fitness for the role of main actor in the spectacle of death, which she envisages as a triumphant didactic spectacle, not one of victimhood. The Revd Watson describes Eliza’s deathbed as “among the most triumphant I have ever witnessed” (1859, 149), thus certifying that her deathbed discourse proved the possibility of personal triumph over death. Taking centre stage for once, this wife and mother found that her Presbyterian faith suddenly raised her up and gave her power to compel others in ways not previously available to her.<sup>12</sup> Eliza’s conscious conversion of her dying into a functional spectacle is matched elsewhere only by the demands

<sup>12</sup> Shaver speaks of dying Methodist women as having been “customarily silenced” (2008, 20) during their lifetimes, and White’s study of the death of Eva in *Uncle Tom’s Cabin* suggests that a curated deathbed provided an opportunity for a child, woman or other marginalized person to gain social and spiritual ascendancy over others (White 1985, 9).

of Stowe's Eva – and then only partly. Eva's role is intended to convert Stowe's readers,<sup>13</sup> while Eliza was unaware that she would ever have readers, and aimed her performance at the immediate spectator group.

Such newfound assertiveness on a deathbed is explained by one scholar as the result of a "conflict between engagement and withdrawal" (White 1985, 12) for women in Victorian culture. Trained to see the domestic role of wife, mother and Sunday-school teacher as her only sphere, a woman like Eliza nevertheless sought validation through her faith. In this more egalitarian sphere, Eliza found fulfilment and vindication. There is even a tinge of vindictiveness in her passive-aggressive call to her neighbour to "see how a Christian can die". Puley stands condemned for not being a true Christian, or not in the sense that Eliza and Joseph Davidson recognized. Though Eliza does not boast, her deathbed account does reveal a sense of social and spiritual triumph (White 1985, 13). Dying well makes Eliza, in her mind, superior to those like Puley who remain living in a condition of dubious salvation. Eliza's good death cannot be fully "self-sacrificial" like the child's deathbed in *Uncle Tom's Cabin* (White 1985, 15). Readers feel an assertive protest on the part of Eliza against some unspecified scepticism about the quality or utility of her faith. Eliza had no official position as the wife of a church elder, but still she had a public role. Not a missionary in life, she becomes one through this performative deathbed, as the Revd Watson affirms: "I improved her death last Sabbath day in *two sermons* ... On both occasions the attendance was large, and a very deep and I hope hallowed impression was left on the minds of us all" (1859, 152).<sup>14</sup> Significantly, obituaries of female missionaries who died abroad in service to their church often essentialized these women (Cassidy 2002, 210), reducing them to the template of acceptable womanly virtues of duty and self-sacrifice. In Eliza's case, however, the Revd Watson has been unable to fit his subject neatly into the template, although he does his best in the article for the *Missionary Record*.

## 6 Conclusion: Comparative Deathbed Spectacles

Table 1 shows that many of the elements of the "good death" are shared by both the fictional and the historical deathbeds.

There is no clear division between Victorian fictional deaths and the three biographical accounts. The Revd Watson wrote Eliza Hitchcock Davidson's deathbed with many of the features of fiction: direct speech, entrances and exits from the death chamber, emotional exclamations, and moral messages. Nevertheless, one account does stand out in Table 1: that of William Carlyle from *East Lynne*. The column of negatives indicates that the goal of a child's deathbed in sensation fiction differed, with no emphasis on the sacrificial death, nor on the deathbed as a catalyst for greater faith. In contrast, the deathbed of Eva from *Uncle Tom's Cabin* exhibits a profile almost identical to that of Eliza Davidson from later in the same decade. Meanwhile, the male pastors check all the boxes except the "angelic". The

<sup>13</sup> Conversion more to the cause of Abolition than to the Christian faith.

<sup>14</sup> By "improve" Watson meant the older sense of "to make my own use of", not that he changed Eliza's discourse. Her good death was successfully instrumentalized to create material for two whole sermons, and thus used to encourage similar piety.

TABLE 1. Deathbeds compared: Y = yes, strongly present; y = yes, somewhat present; n = no, mostly absent; N = no, completely absent

Element	Little Nell 1840/41	Beth March 1868	William Carlyle 1861	Little Eva 1852	M'Cheyne 1848	Cun- ningham 1871	Eliza H. Davidson 1859
Witness(es)	Y	Y	n	Y	Y	Y	Y
Joy, confidence	Y	Y	n	Y	Y	Y	Y
Music, hymns	N	Y	N	Y	Y	Y	Y
Angelic demeanour	Y	y	N	Y	N	N	Y
Advice, exhortations	y	N	N	Y	Y	Y	Y
Scripture readings	N	N	N	Y	Y	Y	Y
Requests, summons	y	y	y	Y	Y	Y	Y

three historical deathbed profiles converge around similar themes. One crucial distinction concerns self-dramatization: M'Cheyne and Cunningham arranged their exits as quiet moments of witness, while Eliza was in no way quiet. Though submissive to her God and her fate, she mounts a spectacle of resisting all doubt about salvation. The account clarifies that Eliza wrung from the experience of dying all the credit available – both sacred and social – making a spectacle to demonstrate “the worth and excellence of personal religion” (Watson 1859, 149). The success of her staging of her own deathbed extends to the existence of this biographical account, which displaced potential narratives of other congregants and more mundane mission activities in the pages of the *United Presbyterian Missionary Record*.

## References

- Alcott, Louisa May. 1915 (1868). *Little Women or Meg, Jo, Beth, and Amy*. Boston: Little Brown & Co.
- Bonar, Revd Andrew A. 1848. *The Life and Remains, Letters, Lectures, and Poems of the Rev. Robert Murray M'Cheyne*. New York: Robert Carter.
- Bross, Kristina. 2001. “Dying saints, vanishing savages: ‘Dying Indian speeches’ in colonial New England literature.” *Early American Literature* 36 (3): 325–52.
- Cassidy, Cheryl. 2002. “Dying in the light: The rhetoric of nineteenth-century female evangelical obituaries.” *Victorian Periodical Review* 35 (3): 206–13.
- Dickens, Charles. 1868. (1841). *The Old Curiosity Shop*. London: Chapman & Hall.
- Doddridge, Philip. 1766. *Hymns Founded on Various Texts in the Holy Scriptures*. London: Job Orton.
- Duncan, Heather. 2018. “Human ‘ish’: Voices from beyond the grave in contemporary narratives.” *ELOPE: English Language Overseas Perspectives and Enquiries* 15 (1): 83–97. <https://doi.org/10.4312/elope.15.1.83-97>.
- Gadpaille, Michelle. 2019. *The Ethical Atlantic: Advocacy Networking and the Slavery Narrative, 1830–1850*. Newcastle upon Tyne: Cambridge Scholars Publishing.

- Hartley, Lodowick. 1949. "The worm and the thorn: A study of Cowper's *Olney Hymns*." *The Journal of Religion* 29 (3): 220–29.
- Jenson, Deborah, and Victoria Szabo. 2011. "Cholera in Haiti and other Caribbean regions, 19th century." *Emerging Infectious Diseases* 17 (11): 2130–35. <https://doi.org/10.3201/eid1711.110958>.
- Lawson, Robert. 1859. "Observations on the outbreak of yellow fever among the troops at Newcastle, Jamaica, in the latter part of 1856." *British Foreign Medico-Chirurgical Review* 24 (48): 445–80.
- Martineau, Harriet. 1844. *Life in the Sickroom: Essays by an Invalid*. London: Edward Moxon.
- Masur, Margo. 2015. "Inhumanly Beautiful": The aesthetics of the nineteenth-century deathbed scene." *English Theses* 18. [https://digitalcommons.buffalostate.edu/english\\_theses/18](https://digitalcommons.buffalostate.edu/english_theses/18).
- Millar, Alexander Hastie. 1900 (1885). "McCheyne, Robert Murray." *Dictionary of National Biography*, Vol. 35, *Wikisource*. Accessed December 30, 2020. [https://en.wikisource.org/wiki/Dictionary\\_of\\_National\\_Biography,\\_1885-1900/McCheyne,\\_Robert\\_Murray](https://en.wikisource.org/wiki/Dictionary_of_National_Biography,_1885-1900/McCheyne,_Robert_Murray).
- Mullan, John. 2009. "Ten of the best deathbed scenes in literature." *The Guardian*, December 5, 2009. <https://www.theguardian.com/books/2009/dec/05/ten-best-deathbed-scenes>.
- Newton, John. 1825. *The Olney Hymns in Three Books*. London: James Duncan.
- Phillippo, James M. 1843. *Jamaica: Its Past and Present State*. Philadelphia: James M. Campbell & Co. / New York: Saxton & Miles.
- Rainy, Robert, and James Mackenzie. 1871. *The Life of William Cunningham, D.D Principal and Professor of Theology and Church History, New College, Edinburgh*. London: T. Nelson & Sons.
- Richardson, Samuel. 1863 (1748) *Clarissa; or, The History of a Young Lady...* Vol IV. Leipzig: Bernhard Tauchnitz.
- Rivett, Sarah. 2011. *The Science of the Soul in Colonial New England*. Chapel Hill, NC: University of North Carolina Press.
- Shaver, Lisa. 2008. "Women's deathbed pulpits: From quiet congregants to iconic ministers." *Rhetoric Review* 27 (1): 20–37. <https://www.jstor.org/stable/20176821>.
- Shelley, Mary. 1994 (1826). *The Last Man*. Edited by Morton D. Paley. Oxford: Oxford University Press.
- Smith, James. 1850. *A Selection of Spiritual Poetry*. London: Simpkin, Marshall & Co.
- Stowe, Harriet Beecher. 1994 (1852). *Uncle Tom's Cabin, or Life among the Lowly*. Edited by Elizabeth Ammons. New York: W. W. Norton.
- Watson, Revd James. 1859. "Kingston. The triumph of faith in Christ." *United Presbyterian Missionary Record*, August 1, 1859, 149–52.
- . 1856. "Proceedings of the United Presbyterian Synod." In *The United Presbyterian Magazine* Vol. X. Edinburgh: William Oliphant and Sons.
- white, isabelle. 1985. "The uses of death in *Uncle Tom's Cabin*." *American Studies* 26 (1): 5–17. <https://doi.org/10.1353/amsj.v26i1.2557>.
- Wood, Ellen. 2000 (1861). *East Lynne*. Edited by Andrew Maunder. Peterborough, ON: Broadview Press.
- Zaaraoui, Karima. 2017. "Death drive, martyrdom and deathbed scenes. The narration of death in Harriet Wilson's autobiographical novel." *E-Rea; Revue Électronique d'Études sur le Monde Anglophone* 15 (1). <https://doi.org/10.4000/erea.6022>.



## Authorship Attribution and the Material Realities of Early Modern Play Texts

### ABSTRACT

The digitization of texts and the advent of big data analyses have transformed our understanding of authorship and collaboration in early modern drama. However, this progress ought to be carefully contextualized within the material realities of early modern playwriting. The scarcity of surviving dramatic manuscripts underscores the significant role of agents like compositors, printers and editors, and the loss of the majority of plays from English commercial theatres casts doubt on the reliability of comparisons based on unique or common verbal parallels. The article focuses on drama from the Elizabethan and Jacobean periods, particularly the recently proposed collaboration between Christopher Marlowe and William Shakespeare on the *Henry VI* plays. Applying the IT concept of GIGO (garbage in, garbage out), it highlights the impact of textual transmission intricacies on authorship attribution, emphasizing that even the most sophisticated attribution techniques are only as reliable as the (often unreliable) data they utilize.

**Keywords:** authorship, Christopher Marlowe, William Shakespeare, textual transmission, material texts, stylometry, GIGO

## Ugotavljanje avtorstva in materialnost zgodnjenovoveških dramskih besedil

### IZVLEČEK

Digitalizacija besedil in analiza velikih količin podatkov močno vplivata na naše razumevanje avtorstva in soavtorstva zgodnjenovoveške dramatike. Ob izjemnem razvoju, ki ga to prinaša, pa je treba upoštevati tudi materialne danosti nastajanja zgodnjenovoveških gledaliških iger. Pomanjkanje ohranjenih dramskih rokopisov nas opozarja na pomembno vlogo posrednikov, kot so stavci, tiskarji in uredniki; ker pa se je izgubila tudi večina dram iz tedanjih angleških komercialnih gledališč, so primerjave, ki temeljijo na odkrivanju edinstvenih ali pogostih besednih zvez pri različnih avtorjih, nezanesljive. Članek se osredotoča na dramatiko iz elizabetinskega in jakobovskega obdobja, zlasti na nedavno predlagano sodelovanje Christopherja Marlowa in Williama Shakespeara pri igrh o Henrik VI. Z vpeljavo koncepta GIGO (garbage in, garbage out) iz informacijske tehnologije opozarja na vpliv, ki ga imajo zapleteni procesi nastajanja in prenosa besedil na avtorstvo, in poudarja, da so tudi najbolj izpopolnjene tehnike ugotavljanja avtorstva zanesljive le toliko, kolikor so zanesljivi (pogosto nezanesljivi) podatki, ki jih uporabljajo.

**Gljučne besede:** avtorstvo, Christopher Marlowe, William Shakespeare, dramatika, elizabetinsko gledališče, stilometrija, GIGO

# 1 Introduction

Literary scholarship does not make global headlines every day. However, the publication of *The New Oxford Shakespeare* (NOS) in 2016 did, and it did so for various reasons, the one most emphasized by the media being the suggestion that Christopher Marlowe had co-authored the three *Henry VI* plays. But when it was announced the news was already more than 200 years old, as there have been dozens of studies examining the authorship of the three *Henry VI* plays, and doubts about William Shakespeare's exclusive authorship, especially of *Part One*, have persisted since the 18th century. Still, the NOS claims to be the first edition of Shakespeare's complete works based on quantitative study rather than editorial taste (Craig and Greatley-Hirsch 2017, 10).<sup>1</sup>

Quantitative authorship studies (also variously known as linguistic processing, computer-assisted methods, stylostatistics, stylometry, stylometrics, computer stylistics, non-traditional attribution methods, etc.) are a fast-developing field of research in authorship. Nonetheless, examining the authorship of early modern English drama poses a formidable challenge, primarily due to the precarious state of extant texts and their ambiguous origins. Accordingly, this article is situated at the intersection of stylometry, the materiality of early modern play texts and their (collaborative) authorship. The proposed collaboration in the NOS between Marlowe and Shakespeare (and others) serves as an illustrative case study of the broader issues concerning the authorship of these works, the lack of solid, conclusive evidence, and the collaborative nature of theatrical and book productions during Elizabethan and Jacobean periods.

## 2 The *Henry VI* Plays

The three *Henry VI* plays – and *Richard III* (together they are sometimes seen as Shakespeare's first historical tetralogy) – chronicle the events surrounding the reign of King Henry VI and the power struggles for the English throne against the backdrop of the Wars of the Roses, between the Yorkists and Lancastrians. The play in the *Henry VI* trilogy (if it is a trilogy at all) that is now generally thought to have been written and staged first was what is now known as *The Second Part of Henry the Sixth* (2 *Henry VI*) after its 1623 Folio text. The earlier version, shorter by a third, was titled *The First Part of the Contention betwixt the Two Famous Houses of York and Lancaster* and first published in quarto in 1594 without an author's name. *The Third Part of Henry the Sixth* (3 *Henry VI*) as it is titled in the First Folio also has a shorter alternative version – the 1595 octavo *The True Tragedy of Richard Duke of York*. Like *The Contention*, *The True Tragedy* was also issued without an author's name. The two plays were first attributed to Shakespeare in Thomas Pavier's 1619 reprint and then more credibly in the 1623 Folio (Taylor and Loughnane 2017, 493–99).

It is now usually thought that *The First Part of Henry the Sixth* (1 *Henry VI*) was written as a prequel to cash in on the successes of *The Contention* and *The True Tragedy*. It exists in only one early edition – the one from the 1623 Folio. There is widespread agreement that the play is collaborative, but no broad consensus has been reached “as to the authors involved or the portions of the play attributed to each of them” (Chernaik 2014, 195). In addition to

---

<sup>1</sup> *The New Oxford Shakespeare* is accessible online at <https://www.oxfordscholarlyeditions.com/nos>.

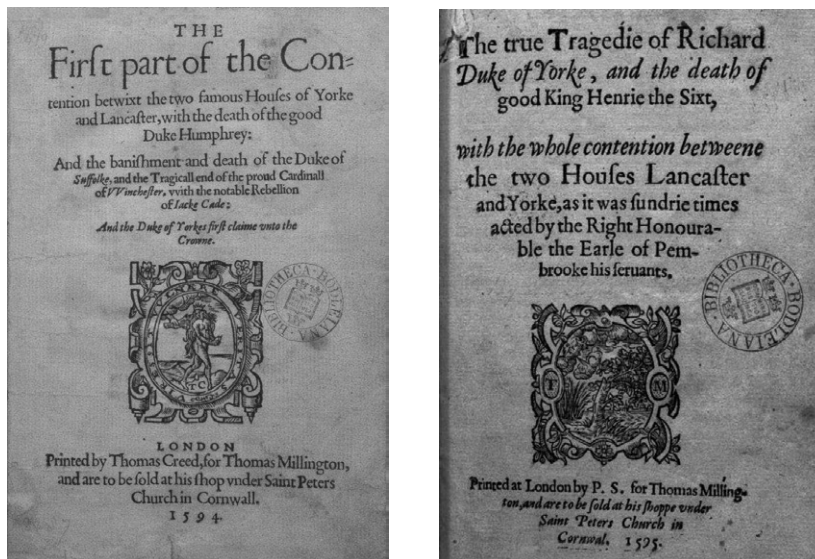


FIGURE 1. The covers of the first editions of 2 and 3 Henry VI, quarto (1594) and octavo (1595) respectively (Title page of the first quarto ... 2011; Title page of the play now known ... 2011).

Shakespeare, the names circulating are Thomas Nashe, Thomas Kyd, Christopher Marlowe, Thomas Lodge and Robert Greene (Taylor and Loughnane 2017, 513–17).

When a play – like *1 Henry VI* – exists in only one early edition, either from the First Folio or in a quarto/octavo edition issued before or later, that is typically taken as the most authoritative version. However, there are often more than one early editions of the same play,

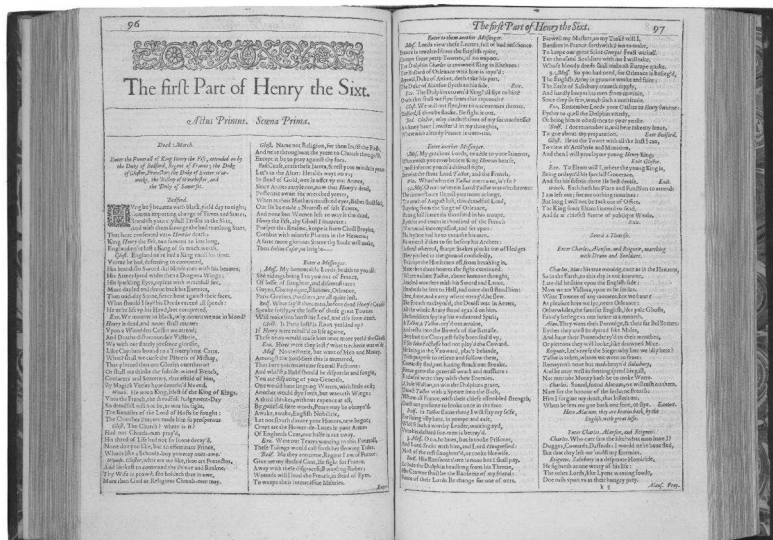


FIGURE 2. The first two pages of *1 Henry VI* in the First Folio (1623) (The first page ... 2015).

often in importantly different variants, as is the case with *2* and *3 Henry VI*. Therefore, are the different texts “versions of the same play” or “different plays written by different dramatists?” (Cox and Rasmussen 2020, 158–59). Were the changes intentional or accidental? Broadly speaking, there are two hypotheses explaining the origins of these textual variants. A very influential idea in the 20th century (promoted by the New Bibliography and especially W.W. Greg and A.W. Pollard) was that the quarto/octavo texts were memorial reconstructions of the texts behind the Folio texts. The advocates of this, now contested hypothesis argue that actors reconstructed the text of a popular play from memory (or through note-taking) in order to sell it on to a publisher or perform it when touring the provinces (Murphy 2007, 8). On the other hand, the *NOS* editors of the *Henry VI* plays champion the revision theory, which proposes that the quarto/octavo variants were written in collaboration and later revised or adapted by Shakespeare to produce the Folio texts (Taylor and Loughnane 2017, 493–99). In the *NOS*, *2 Henry VI* is attributed to Shakespeare, Marlowe, and another unidentified author and thought to have been revised by Shakespeare (original best guess 1590, revision best guess 1595); *3 Henry VI* to Shakespeare, Marlowe, and another unidentified author, with revisions by Shakespeare (original best guess late 1590, revision best guess 1595);<sup>2</sup> and *1 Henry VI* to Nashe, Marlowe, and another unidentified author, adapted by Shakespeare (original date March 1592, adaptation best guess 1595).<sup>3</sup> According to the *NOS*, Shakespeare was “the dominant writer” in the three *Henry VI* plays and *Richard III*, “the only one present in all four plays, and the reviser who turned them into a coherent historical and dramatic sequence” (Taylor and Loughnane 2017, 425).

### 3 Stylometry

Different sorts of evidence are considered when attributing the authorship of early modern English drama (e.g., palaeographical evidence, the theatrical provenance of individual play texts, chronological evidence, poetic and linguistic evidence, and others (Taylor and Loughnane 2017)). Whereas traditional attribution was based on “little more than personal poetical taste” (Egan 2017, 28), the methods evolved and became ever more sophisticated, ranging from verse and metrical tests (rhymes, number of feet, incomplete lines, enjambement, pauses), analysis of unique occurrences of words, spelling, function words (prepositions, conjunctions, articles, particles, auxiliary verbs, and pronouns), rare and/or common vocabulary, sentence length measured in words, word length, verbal parallels (n-grams), and image clusters (Egan 2017, 29–42) to 21st century stylometric tests based on computer-assisted data analyses (e.g., Delta, Nearest Shrunken Centroid, Random Forests, Zeta, PCA, Shannon entropy, and the *t*-test (Craig and Greatley-Hirsch 2017)).

Gabriel Egan (2017, 33) notes that until the second half of the 20th century there were, in effect, only two methods in use: counting verse features and looking for verbal parallels. This changed fundamentally in the early 21st century, which has seen “an explosion of methods

<sup>2</sup> By contrast, Darren Freebury-Jones (2016; 2017; 2022) insists on assigning *2* and *3 Henry VI* to Shakespeare alone. Moreover, while he supports the common view on “the pervasive influence that Marlowe had on Shakespeare’s dramaturgy”, he finds “no compelling evidence to suggest that Marlowe and Shakespeare co-authored any surviving play texts” (Freebury-Jones 2022, 181, 182).

<sup>3</sup> Others, e.g., Brian Vickers (2008) and Freebury-Jones (2022), ascribe *1 Henry VI* to Kyd (and Nashe) and Shakespeare.

for attributing sections of early modern drama” (Gossett 2022, 93). As Heather Hirschfeld (2015, 23) writes, the emergence of machine-readable databases, including Elizabethan and Jacobean plays, and computerized analyses and comparisons of linguistic features “have enabled a kind of ‘renaissance’ in attribution studies”. However, although the methods claim to be scientific and to have “revolutionized attribution studies” (Jackson 2017, 49), they have so far failed to yield generally accepted, sufficiently convincing results. Typically, they do not – cannot – give a “yes” or “no” answer. Rather, they employ assumption, conjecture, and statistical probabilities.

Describing authorship attribution endeavours at the beginning of the 20th century, Samuel Schoenbaum (1966, 68) observed that “the metrical tables, the word-lists, the impressionistic comments on theme, character, and plot, the final brandishing of the parallels – all designed to create a ‘cumulative impression’ – were to constitute a durable formula”. Still today circumstantial evidence is occasionally presented by some scholars in a way that relies on the aggregate weight of individual pieces of evidence with limited value. Presenting evidence in such a way, these scholars hope that their collective strength might outweigh the weaknesses of each individual piece (Love 2002, 203). These methods in general, and the *NOS*’s in particular, have been subject to much criticism and controversy, both in terms of their methods and their findings. Different authors have emphasized that there is inadequate theoretical framework to explain the successes of the methods used (Evert et al. 2017), that there may be some academic vested interests behind the *NOS*’s attributions (Williams 2018), and that although the edition prides itself on “big data” analyses, some of the samples studied are very small (Auerbach 2019; Barber 2021). To quote a reviewer: the collaborations newly determined by the *NOS* “should be taken *cum multo salis*” (Rudman 2019, 705).

On the other hand, there is a great deal of optimism among the proponents of these new stylometric methods, who argue that “given data enough and time to develop the necessary methods, quantitative criticism promises to shed light on many unresolved questions about Shakespeare’s canon, chronology, sources and style, as well as his relationship to his contemporaries” (Greatley-Hirsch 2021, 215). They contend that, contrary to stylometric findings, subjective preferences can lead individuals to reject even the most compelling authorship attribution that contradicts their personal beliefs (Jowett 2007, 19). They maintain that the field of authorship attribution has progressed to the point where subjective impressions are no longer considered sufficient to support new attributions, and quantitative studies are now necessary to provide a more rigorous and persuasive basis for attributing works to specific authors (Craig and Greatley-Hirsch 2017, 9). It also seems sensible to accept that “the fact that a technique is badly applied in one case does not mean it may not be well applied elsewhere” (Love 2002, 79). Nevertheless, it is just as important to bear in mind that “using modern digital technology does not automatically immunize scholars against perpetuating old methodological mistakes” (Jackson 2017, 49).

The argument that I would like to develop further in this article does not dispute the techniques themselves, since most serious analyses consistently outperform subjective approaches and chance discoveries, and demonstrate their ability to extract meaningful insights from the data. While some authorship attribution methods certainly seem more

reliable (“scientific”) than others, I would like to argue that all the methods and techniques have the same fundamental problem – and that it precedes them. I propose that what really matters is the material realities behind the data – that is, the origins of the Elizabethan and Jacobean play texts under investigation.

## 4 GIGO

At this point, I would like to introduce a concept from computer science – GIGO (garbage in, garbage out), which means that the quality of output is intrinsically linked to the quality of input.<sup>4</sup> In other words, if inaccurate or flawed data is fed into a system, the resulting output is unlikely to be reliable or meaningful. To state the obvious: the data – the “input” – in authorship attribution studies is the text; thus, the GIGO principle underscores the pivotal correlation between the origin, nature, and quality of texts analysed and the subsequent dependability of outcomes aimed at discerning their authorship. For example, if the texts used for stylistic analysis are of poor quality, corrupted, not representative of the author’s style because of their transcription or transmission or if they are misattributed, the results will inevitably be compromised. Similarly, if there are biases in the training data or erroneous assumptions made during the study, they can propagate through the analysis, leading to questionable results. As Egan (2017, 27) observes, “Shakespearean authorship attribution by computational stylistics [...] has already had spectacular failures because the value of evidence was wrongly weighed”.

In an early awareness of GIGO in attribution studies, the Victorian scholar Frederick G. Fleay, when examining *hapax legomena* in Shakespeare’s plays and relying for that on Mary Cowden Clarke’s concordance, acknowledged that any potential errors in her work may have caused misattribution in his (1874, qtd. in Egan 2017, 30).<sup>5</sup> In an example from 1901, Ashley Horace Thorndike used the distinction between the pronoun “them” and its contracted form “em” to distinguish between Beaumont, Fletcher, and Massinger. He found that Massinger was the only one to have invariably used the longer form throughout his texts – until he realized he had been using an edition of Massinger’s texts whose editor changed all occurrences of “em” to “them” (Schoenbaum 1966, 171). The researchers who continue to study authors’ preferences, say, between “whilst” and “while”, “among” and “amongst”, “betwixt” and “between” (Freebury-Jones 2019, 4) will certainly be aware of any possible printing-house or editorial interventions in the texts they use for their analyses. As mentioned above, some authorship attribution studies look at sentence length. However, there may be significant differences in how the same text is divided into sentences in different version (e.g., *Hamlet* in Q1 (1603), Q2 (1604/5), and F (1623) (Murphy 2007, 5)); therefore, the choice of the edition(s) the scholar chooses to work with is, again, of key importance. Another example of editorial intervention with wide-ranging consequences is “distinguishing between verse and prose, and affirming line endings”, which are “as subjective and delicate tasks as

<sup>4</sup> While online search engines and the *Oxford English Dictionary* (s.v. “garbage”) suggest that the term dates to the 1950s, the concept itself was already known to computing pioneers like Charles Babbage (1791–1871), who invented, *inter alia*, early mechanical computers. He conveyed his astonishment at the “confusion of ideas” expressed to him by members of the Upper and Lower Houses, asking him “if you put into the machine wrong figures, will the right answers come out?” (Babbage 1864, 67).

<sup>5</sup> Incidentally, Fleay believed all the three *Henry VI* plays to have been “wrongly ascribed to Shakespeare” (1874, qtd. in Egan 2017, 30).

any in editing” (Jowett 2007, 137). During analyses, including attribution, tabulating verse and prose – and deciding what paratextual features to include in the counting – “is a highly inexact process”, and “scholars not infrequently disagree as to what is prose and what is verse in the texts they edit” (Bruster 2023, 136).

A prominent scholar in stylometry, Hugh Craig (2021, 230, 232), who asserts “that there are persistent internal consistencies in authorial canons”, adds himself that occasionally “authors can confound standard authorship methods and write unlike themselves throughout a work”. In an example from the work Craig and John Burrows did while testing for the authorship of *3 Henry VI* for the NOS, they describe their stylometric tests for Marlowe’s *The Jew of Malta* as “a spectacular failure, with only 29 out of the 81 segments attributed to Marlowe” (Burrows and Craig 2017, 210). Their explanation is that the play “evidently departs from the Marlowe style as represented by the other six plays”, but it is nevertheless “still Marlovian” (Craig 2017, 212). Yet two other renowned Shakespeare scholars, writing for the same volume, draw a conclusion that challenges that interpretation, stating that one cannot use “the same scrupulous methods” in different cases and only accept their conclusions in some cases and not in others: “once the validity of internal evidence is granted in even a single case, then its application to other cases cannot be avoided” (Taylor and Loughnane 2017, 431, 432).

The corpus sizes of the dramatists considered for authorship present additional difficulties, especially when compared to Shakespeare’s relatively large and unambiguous corpus. Thomas Kyd, for instance, only has one well-attributed original play to his name, *The Spanish Tragedy*, so authorship scholars may add his translation *Cornelia* and an anonymous play, *Soliman and Perseda*, to his corpus (see, e.g., Craig 2021, 235). This, of course, is speculative and prone to returning doubtful results. Thomas Nashe similarly has only one sole-authored play, the comedy *Summer’s Last Will and Testament*, which is why some studies exclude him from the pool of the possible authors they survey for the authorship of *1 Henry VI*. On the other hand, the editors of the NOS do not: “Although we cannot be certain that Nashe wrote every line, scene, or sub-scene of Act 1, he is clearly the primary author of Act 1” (Taylor and Loughnane 2017, 514).

The application of GIGO in early modern English drama authorship studies underlines the importance of rigorous and careful data collection, textual analysis, and statistical modelling to ensure that the conclusions drawn from these studies are trustworthy and reflective of the actual authorship patterns. To quote Taylor and Loughnane (2017, 432), “the issue is not whether to use stylistic evidence at all, but how to use it properly, how to distinguish strong stylistic evidence from weak stylistic evidence”. Indeed, “textual scholars past and present differ absolutely on some of the same evidence” (Knowles 2001, 139). This, in turn, leads to questions about what defines strong evidence and how to interpret it.

## 5 Texts and Authorship

It is far from straightforward to identify unequivocally what William Shakespeare or Christopher Marlowe (or most of their contemporaries) wrote in their own hand, and to what extent the printed texts attributed to them represent it. John Jowett (2007, 18) argues that “documentary evidence such as title pages” should take preference over “‘internal’ evidence

based on language and stylistic preferences”. However, external evidence is often unreliable. There are at least three problems related to author attribution on Elizabethan and Jacobean title pages. First, authors working for the commercial theatre in the late 16th and early 17th centuries had very little control over their plays (Jowett 2007, 7), which – as material objects (manuscripts) – became “the legal property” of the theatre company that bought them (Taylor and Loughnane 2017, 422; Giddens 2011, 13). Shakespeare’s name first appeared on title pages in 1598, although some of his plays continued to appear anonymously after that, too. Marlowe’s name first appeared on a title page in 1594. Second, the authenticity of external evidence for plays’ authorship is often subject to debate, as publishers may have made errors or acted fraudulently in attributing works to specific authors (Giddens 2011, 60; Elliott and Greatley-Hirsch 2017, 141). Third, the “title pages of plays from the early modern public theatres drastically under-represent collaboration” (Jowett 2007, 18).

Shakespeare’s canon was essentially defined by the publication of his collected plays, *Mr. William Shakespeare’s Comedies, Histories & Tragedies*, known as the First Folio, in 1623. This remains “the chief external evidence” for his works (Taylor and Loughnane 2017, 417). Nevertheless, despite the later inclusion of *Pericles* and *The Two Noble Kinsmen*, both of which are absent from the First Folio, in the playwright’s oeuvre, “there are undoubtedly more plays to which Shakespeare contributed parts, and substantial parts of plays in the 1623 First Folio are not his” (Egan 2017, 28). While Shakespeare’s First Folio was published by his colleagues, albeit seven years after he had died, Marlowe’s first collected works was only published 230 years after his death, in 1823 (Hopkins 2005, 184). On the other hand, and unlike Shakespeare’s early editions, all Marlowe’s plays were published with his name on the title page – except for the two parts of *Tamburlaine the Great* (1590), which were, in contrast, the only play texts of his to be published before his death in 1593. But in fact Marlowe’s authorship is much more complicated than this summary implies. Indeed, the *Tamburlaine* plays remained unattributed for decades, and “by the most conservative standards” they should be regarded as anonymous (Erne 2013, 64). *Dido, Queen of Carthage* (1594) was attributed on the title page to Marlowe and Thomas Nashe, and this “attribution to two authors in 1594 is not just exceptional but unique, coming at a time when few plays were attributed to one author” (Hadfield 2021, 205). *The Massacre at Paris* (1594?) is typically seen as a problematic text, as it “has evidently suffered in transmission” (Craig 2021, 235). *Doctor Faustus* exists in two strikingly different early editions, the A-text of 1604 and the B-text of 1616, and they are both collaborative play texts. The first (extant) text of *The Jew of Malta* was not published until 1633, by which time it had likely been revised by others (Craig 2021, 235). Thus, when it comes to Marlowe only *Edward the Second* (1594) comes anywhere near the definition of a non-contentious, sole-authored, well-attributed play text required for unambiguous attribution tests.

Expanding upon the earlier assertion, the input data in authorship attribution studies is texts, and regardless of how accurate an attribution technique is, the results will be inaccurate if the input is unsound. This is where serious attribution troubles begin. Given the realities of early modern play text production, this is true for sole-authored and collaborative plays, for well-attributed and anonymous ones (as well as all those in the grey areas in between). The stark fact is that there are no undisputed manuscripts of any of Marlowe’s or Shakespeare’s

undisputed works in their own handwriting, and no completely reliable early definition of their canons (Taylor et al. 2017, v). All we have that is relatively reliably in Shakespeare's and Marlowe's hands is a few signatures on legal documents (six by the former and one by the latter). Other than that – and a handful of documents, comments by fellow authors and anecdotal evidence from their contemporaries – the two authors are defined by the plays and poems that are now associated with their names (Gossett 2022, 1).

Authorship studies must face another major obstacle: “there is no perfect early printed play” (Giddens 2011, 116). In general, when a play by an early modern playwright, say Marlowe or Shakespeare, was set into type from a manuscript, there were an array of possibilities for the printer's copy: a manuscript in the author's hand, a copy of that manuscript (prepared by a scribe, perhaps one approved by the Master of the Revels), a copy prepared for the bookholder (with stage managing notations) or an amalgam of a number of different manuscripts (Berger 2007, 62). Of course, printers also worked from previously printed texts. Whatever the compositor used as his copy may have been marked up, amended, annotated, revised, cut, added to, or otherwise changed by a variety of textual transmission agents (e.g., actors, scribes, censors, or authors themselves) (Murphy 2007, 11). Consequently, any attempt to reconstruct what Shakespeare or Marlowe really wrote “requires a degree of scholarly reconstruction, and an acceptance that much of this reconstruction is likely to remain hypothetical” (Shrank and Werstine 2021, 67).

In her examination of early modern English plays, Tiffany Stern (2009, 1) finds that they were “pieced together out of a collection of odds and ends”, such as plots drawn up by professional plotters, scenes written out of sequence, prologues and epilogues, songs, scrolls, arguments, and playbills. Each element had its own origin (i.e., author), circulation, and rate of loss and survival (Stern 2009, 2–4). In this way, commercial theatres helped create an explosion in the production of “licensed playbooks, actors' parts, musicians' scores, prologues, epilogues, scenarios, backstage plots, perhaps scrolls and letters to be read on stage, scribal copies for fans and patrons (and perhaps booksellers), account books, inventories of costumes and properties, contracts, and advertising copy for playbills” (Taylor 2017, 8).

Elizabethan and Jacobean dramatists participated (and collaborated) in a network of playwrights, revisers, scribes, printers, compositors, correctors, actors, censors, editors, and theatrical impresarios responsible for shaping early modern plays. Accordingly, books, plays – even authors themselves – are sometimes seen “as unstable entities produced within intersecting cultural and commercial networks” (Hooks 2016, 157). This complex chain of textual production can make it very hard to talk about authors and authorship in our modern sense. It is thus crucial to understand that all these agents had their own agendas (Taylor 2016, 143) and “produced the plural Shakespeares” (Stallybrass and Chartier 2007, 35).<sup>6</sup>

<sup>6</sup> Several prominent scholars dispute such emphasis on non-authorial agents, arguing that the author's style trumps secondary factors like printers and compositors (Egan 2014), which “have minimal effect on the analysis of large numbers of data” (Jackson 2017, 124). They maintain that “within collaborative texts it is possible to tell writers apart; even if we can't identify them, we can nonetheless frequently posit shifts in authorial hands through local differentiation across a range of stylistic indicators” (Sharpe 2014, 33).

## 6 Collaborations

A substantial body of scholarly investigation now supports the notion that Shakespeare – like many of his contemporaries – engaged in collaborative writing. Consequently, the canons of early modern playwrights are now not only unstable and open, but also intersectional, as the same works make part of the canons of different authors: “We now know that any ‘Complete Works’ of Shakespeare is also an anthology of selected work by Middleton, Fletcher, Marlowe, Peele, Nashe, Wilkins, and the Anonymous” (Taylor 2016, 148).<sup>7</sup> However, the statistics on the shares of single-authored versus collaborative plays (and within the latter the lines, scenes, and acts written by individual authors) are hard to pin down. Joseph Rudman (2016, 315) describes 16 out of the 36 First Folio plays “as being collaborations or at least having significant interpolations”. Gary Taylor (2016, 141) suggests that “anywhere from a quarter to a third of Shakespeare’s plays contain material written by other professional playwrights. And even those numbers underestimate the pervasiveness of collaboration in the Shakespeare canon”. The authors of the *NOS* “identify Shakespeare’s as the only hand in fewer than two-thirds of the surviving plays that Shakespeare had a hand in” (Taylor 2017, 23). On the other hand, they believe that he wrote “at least 90 per cent of the words included in the Folio” (the rest having been contributed by Fletcher, Middleton, Marlowe, Nashe and Peele) (Taylor and Loughnane 2017, 425). The same scholars also argue that “Shakespeare collaborated less frequently” than some of his contemporaries (e.g., Fletcher, Beaumont, Middleton, or Massinger) and that “he was unusually independent and solitary” (Taylor and Loughnane 2017, 423). In the same way, Shakespeare is described as “probably unusual in the extent to which he wrote as sole author” (Jowett 2007, 17) and, in terms of attribution challenges, “a haven of safety with twenty-eight well-attributed single-author plays” (Craig 2021, 236).

With the “resistance to the mono-authorship monopoly” (Taylor 2016, 142) came the need to identify individual authors in collaborative texts. Attribution studies require well-attributed, sole-authored plays, which make possible the creation of individual authors’ style markers through stylometric tools. “In an ideal universe,” Elliott and Greatley-Hirsch (2017, 141) write, the studied authors’ corpora “would consist only of well-attributed, sole-authored texts of sound provenance, with each of the individual authors represented by equally sized bodies of writing”. As, by definition, such an ideal universe does not exist, the practical and material realities of the production of early modern play texts must be considered.

There is little doubt today that Shakespeare’s plays were shaped by “several varieties of collaborative authorship” (Giddens 2011, 10). However, there is no universal agreement as to what plays or what parts of the plays are collaborations or who the collaborators were. A significant share of what is commonly attributed to Shakespeare is, in fact, the result of collaborative efforts or adaptations of existing works by other individuals. While writing a play for the commercial stage a playwright could work alone or with others, but he also made use of others’ plays, stories, poetry, chronicles, proverbs, and so on. In Taylor’s (2016, 145) words, “Shakespeare made an honest living stealing other men’s work”. Hence the title pages of, for instance, *Julius Caesar* and *Anthony and Cleopatra* should read “based on Plutarch’s

<sup>7</sup> See, e.g., *Oxford Scholarly Editions Online* ([oxfordscholarlyeditions.com](http://oxfordscholarlyeditions.com)), where *1 Henry VI* appears among the works of Shakespeare, Marlowe, and Nashe.

*Lives*, translated by Thomas North, adapted by Shakespeare” (Taylor 2016, 144). Even if “the actual penning of speeches and scenes was done separately”, which “makes it much easier to differentiate the work of two (or more) authors of a cowritten play” (Taylor 2016, 146), the various possible agents of collaboration pose significant challenges to authorship attribution.

Actors and theatre companies represent another agent of collaboration. Tellingly, until the late 1590s playwrights’ names were frequently left out from plays’ title pages, but companies were “almost always identified” (Jowett 2007, 62). Information about performance was typically given where today’s readers expect the author’s name (Smith 2021, 99–100). That is why some scholars argue for authorship attributions like “Harry the Sixth, as presented by Lord Strange’s Men” (Burns 2000, 84) and for making sense of Elizabethan and Jacobean play production by compiling collected plays produced by the companies different playwrights worked for (e.g., “*The King’s Men’s Complete Works*” (Boguszak 2017, 312)). These authors maintain “that playing companies developed distinctive and recognisable styles – relocating the Foucauldian ‘author function’ from the playwright to the company and its agents” (Craig and Greatley-Hirsch 2017, 166), to which empirical stylometry is opposed in theory and practice, arguing that we should follow “the fundamental ethical principle of giving people credit for the work that they have done” (Taylor 2017, 20; see also Egan 2014; Craig and Greatley-Hirsch 2017).<sup>8</sup>

Another collaborator the playwright “had no option but to collaborate with” was the Master of the Revels’ office, who licensed plays for public performance and may have censored them by, for example, removing expletives or politically sensitive content (Murphy 2007, 14). The final censoring and corrections were typically done “before the play was returned to the playhouse and the actors’ parts were created” (Stern 2004, 145). This is another element of early modern play text production that modern editors occasionally seek to undo (Jowett 2007, 116), but they can only do so in a subjective and speculative manner.

Our only access to Shakespeare’s and Marlowe’s texts is through printed books, and “it is an axiom of book history that authors do not make books. Books are material objects created by specialist craftsmen” (Taylor 2016, 143). As a text was being prepared for printing, its spelling and punctuation “were primarily determined by the compositor, who need not in any case have been working from an authorial manuscript” (Love 2002, 196). Compositors regularly applied their own orthographic conventions to the manuscripts they set in type (Jowett 2021, 88). Thus, the compositors’ role extended beyond simple mechanics; as they (re-)shaped the text, they produced “the material precondition of all later readings” (Stallybrass and Chartier 2007, 37). Furthermore, printers’ in-house proofreading was done while the presses were running (the so-called stop-press corrections), which resulted in a random patchwork of corrected and uncorrected pages in surviving copies (Smith 2021, 95). This occasioned variants not only across different editions but also within the same edition, and these variants are not apparent to the reader who consults only one copy (perhaps the one available in a digitalized copy online). The “wide variations in orthography and typography and poor

<sup>8</sup> While individual dramatists should undoubtedly be credited for their work, the texts “we read and interpret are objects produced by anonymous compositors, pressmen, and proofreaders” (Taylor 2016, 143), whose work ought to be taken into account, too.

proofreading of early printed texts” make the attribution attempts that look at “a mechanical solution to authorship based on style [...] inconclusive at best and widely inaccurate at worst” (Cox and Rasmussen 2020, 47).

Just as we have no undisputed manuscript written by Marlowe or Shakespeare, we can never have any unmediated access to their imaginations. Every edition of the plays they wrote, “from the first that was published to the most recent, has been edited: it has come into print by means of a tangled social process and inevitably exists at some remove from the author” (Greenblatt 1997, 71). All their texts “are editions; all have been edited; all have been mediated by agents other than the author” (Taylor 1987a, 1). Any unmediated text is therefore “unattainable; we can only choose which mediator(s) to accept” (Taylor 1987a, 3). When Richard Jones first published the *Tamburlaine* plays in 1590, he acknowledged in his prefatory epistle “To the gentlemen readers and others that take pleasure in reading histories” that he had “(purposely) omitted and left out some fond and frivolous gestures” which he deemed “unmeet for the matter” and “tedious unto the wise” (qtd. in McInnis 2020, 2). This means that we only have access to Jones’s version and “what *Tamburlaine* would have looked like had Jones not intervened remains a mystery” (McInnis 2020, 2). Meanwhile, in their address “To the great variety of readers” in the First Folio, John Heminges and Henry Condell – Shakespeare’s friends and compilers of the 1623 complete edition of his plays – assure their readers they would no longer be “abused with divers stolen and surreptitious copies, maimed and deformed”, as the texts they provide in their edition are “cured and perfect of their limbs” (qtd. in Gossett 2022, 42). The list of modern editors’ interventions is extensive (including “correction of printing errors, emended readings, relineation, repunctuation, reassignment of speeches, correction and supplying of stage directions, act and scene divisions, the identification of locales, etc.” (Sanders 1990, 20)) – and, as I have argued throughout this essay, they can have significant impact on our views of authorship and style.

## 7 Lost Plays

Sometimes early copies are lost and sometimes they are found. For instance, the 1595 octavo edition of *3 Henry VI* (i.e., *The True Tragedy*) was unknown until 1796 (Cox and Rasmussen 2020, 159). Or, an example from Marlowe: *Edward the Second*’s Q1 (1594) is an exceptional rarity today, and since the Second World War, during which the Kassel copy disappeared, it was believed that the only surviving copy was held in the Zentralbibliothek Zürich. In 2012, however, Jeffrey Masten discovered a third known copy (i.e., the second copy still extant today) in the Universitätsbibliothek Erlangen-Nürnberg. None of these three copies is completely identical to the others (McInnis 2015, 310). Therefore, if one were being pedantic, even non-contentious, well-attributed, sole-authored plays like *Edward the Second* should be attributed in the following manner: the Marlowe of the Q1 *Edward the Second* (or even the Marlowe of the Zurich Q1 *Edward the Second*).

It is estimated that in the lifetime of English commercial playhouses, roughly between 1567 and 1642, about 3,000 plays were written and staged, of which 543 are still extant, 744 are known to have existed but are now lost and there are hundreds that have disappeared without

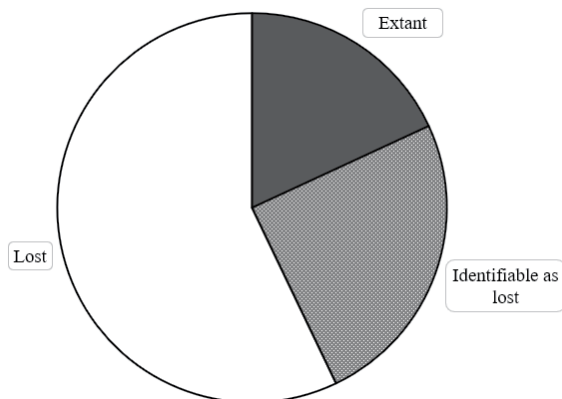


FIGURE 3. Plays from commercial theatres, 1567–1642 (data from McNinnis and Steggle 2014, 1).

a trace (McInnis and Steggle 2014, 1).<sup>9</sup> These figures, as (in)accurate as they may be, have a direct and highly significant bearing on authorship analyses.

Play texts became lost for numerous different reasons, and “the surviving drama is, statistically speaking, atypical precisely because of its survival; these plays constitute the distinct minority of the total dramatic output for the period” (McInnis 2021, 3). The texts that survived did so because they were “able to make the transition into the early modern print marketplace. Those qualities that made a play viable in that marketplace will be over-represented in the survivors” (McInnis and Steggle 2014, 2). Hence, the surviving versions of plays do not necessarily represent the most recent or authoritative versions (McInnis 2021, 7).

Authorship attribution scholars often look at n-grams (consecutive word sequences) or verbal parallels, phrases, and collocations that authors share and then decide (often subjectively) whether the specific intertextual element is a sign of authorship, allusion, imitation, citation, and suchlike – or mere coincidence. Using modern electronic corpora and digital databases, they claim to be able to “ascertain just how many times a verbal repetition occurs in texts of the period” (Freebury-Jones 2021, 191) and determine whether the phrases are “rare (and therefore potentially evidential) or common (and therefore meaningless)” (Taylor and Loughnane 2017, 448). But, to quote Harold Love (2002, 151), “what is singularly lacking to date is any way of calculating the level of significance that applies to such findings”. The crucially important qualifications are the following: first, we can only computer analyse the texts that have been digitalized; and second, because of the enormous gap created by the lost plays we cannot perform adequate negative checks to establish that “similar results are obtained by no other candidate [...] when the same processes are carried out” (Jackson 2017, 59). Clearly, as Gary Taylor (1987b, 89) puts it, “when evidence runs out, opinion runs in”.

<sup>9</sup> This is especially true of the 1580s and early 1590s, which is why “Shakespeare’s so-called Lost Years (1586–91) may have partly been spent writing what are now lost plays” (Taylor and Loughnane 2017, 422).

## 8 Conclusion

Scholars engaged in the examination of individual texts are well-acquainted with the inherent ambiguities within each work, but when extensive datasets are input into computational systems the intricacies and nuanced details may become obscured or overlooked. One thing is certain, though: no text by Shakespeare or Marlowe (or any other early modern playwright) we read today, and in whatever medium we do so, can ever fully represent what they actually wrote in their hands. In assessing the Marlowe-Shakespeare collaboration hypothesis in the *Henry VI* plays – just as in any other authorship attribution study – it is crucial for scholars to evaluate rigorously the quality of historical evidence, the integrity of the texts, the appropriateness of their analytical methods and the potential biases in the interpretation of data. There is no avoiding the fact that “the research findings are only as good as the databases and their users” (Freebury-Jones 2019, 7). Furthermore, the substantial number of lost plays makes negative checks impossible, and leaves gaps in the historical record that render definitive authorial comparisons and determinations elusive.

The choice of a text variant will impact any analysis we perform on it, including authorship attribution. Even when a text is largely uncontroversial, it will incorporate work by agents other than the playwright, as each play is a unique amalgam of original authorial writing (either single-authored or collaborative), theatrical adaptation, printing-house idiosyncrasies, censorship, cutting, revision, adaptation, and editorial mediation. These “historically determined collaborations of authorial and non-authorial intentions” can help explain at least some conflicting findings and attributions, and despite the great promise of empirical stylometry, “recent developments suggest that the question of authorship has not been resolved because it is unresolvable” (Cox and Rasmussen 2020, 47).

To reiterate the essence of my argument, I propose that the reason why different researchers cannot reach any overwhelming consensus or why they attribute the same textual sites to different authors lies not (necessarily) in the methods they use (although these, too, have been subject to rigorous criticism), but rather in the ambiguous origins of the texts they study. In the future, authorship attribution techniques may progress beyond what we deem possible today (very likely with the assistance of AI), but for now the existing textual witnesses (in our case, Elizabethan and Jacobean play texts) do not seem to warrant the degrees of certainty that scholars working in authorship attribution often express. Although this may seem defeatist, I would like to suggest that there is some value in highlighting what we do not know, what is perhaps unknowable, as well as in avoiding unfounded speculation.

## References

- Auerbach, David Benjamin. 2019. “Statistical infelicities in *The New Oxford Shakespeare Authorship Companion*.” *ANQ: A Quarterly Journal of Short Articles, Notes and Reviews* 33 (1): 28–31. <https://doi.org/10.1080/0895769X.2018.1559023>.
- Babbage, Charles. 1864. *Passages from the Life of a Philosopher*. London: Longman.
- Barber, Ros. 2021. “Big data or not enough? Zeta test reliability and the attribution of *Henry VI*.” *Digital Scholarship in the Humanities* 36 (3): 542–64. <https://doi.org/10.1093/llc/fqaa041>.
- Berger, Thomas L. 2007. “Shakespeare writ small: Early single editions of Shakespeare’s plays.” In *A Concise Companion to Shakespeare and the Text*, edited by Andrew Murphy, 57–70. Malden: Blackwell.

- Boguszak, Jakub. 2017. "“Scarlet Experiment!”: *The New Oxford Shakespeare* and the Importance of Authorship." *Shakespeare* 13 (4): 309–12. <https://doi.org/10.1080/17450918.2017.1406394>.
- Bruster, Douglas. 2023. *Seeing Shakespeare's Style*. New York: Routledge.
- Burns, Edward. 2000. "Introduction." In *King Henry VI: Part 1* by William Shakespeare, edited by Edward Burns, 1–103. London: The Arden Shakespeare.
- Burrows, John, and Hugh Craig. 2017. "The Joker in the pack? Marlowe, Kyd, and the co-authorship of *Henry VI, Part 3*." In *The New Oxford Shakespeare: Authorship Companion*, edited by Gary Taylor and Gabriel Egan, 194–217. Oxford: Oxford University Press.
- Chernaik, Warren. 2014. "Shakespeare as co-author: The case of *1 Henry VI*." *Medieval & Renaissance Drama in England* 27: 192–220. <http://www.jstor.org/stable/24322642>.
- Cox, John D., and Eric Rasmussen. 2020. "Introduction." In *King Henry VI: Part 3* by William Shakespeare, edited by John D. Cox and Eric Rasmussen, 1–176. London: The Arden Shakespeare.
- Craig, Hugh. 2021. "Shakespeare and authorship attribution methodologies." In *The Arden Research Handbook of Shakespeare and Textual Studies*, edited by Lukas Erne, 225–43. London: The Arden Shakespeare.
- Craig, Hugh, and Brett Greatley-Hirsch. 2017. *Style, Computers, and Early Modern Drama: Beyond Authorship*. Cambridge: Cambridge University Press.
- Egan, Gabriel. 2014. "What is not collaborative about early modern drama in performance and print?" *Shakespeare Survey* 67 (2014): 18–28. <https://doi.org/10.1017/SSO978110775572.002>.
- . 2017. "A history of Shakespearean authorship attribution." In *The New Oxford Shakespeare: Authorship Companion*, edited by Gary Taylor and Gabriel Egan, 27–47. Oxford: Oxford University Press.
- Elliott, Jack, and Brett Greatley-Hirsch. 2017. "Arden of Faversham, Shakespearean authorship, and 'The Print of Many'." In *The New Oxford Shakespeare: Authorship Companion*, edited by Gary Taylor and Gabriel Egan, 139–81. Oxford: Oxford University Press.
- Erne, Lukas. 2013. *Shakespeare as Literary Dramatist*. 2nd ed. Cambridge: Cambridge University Press.
- Evert, Stefan, Thomas Proisl, Fotis Jannidis, Isabella Reger, Steffen Pielström, Christof Schöch, and Thorsten Vitt. 2017. "Understanding and explaining delta measures for authorship attribution." *Digital Scholarship in the Humanities* 32, suppl. 2: ii4–ii16. <https://doi.org/10.1093/llc/fqx023>.
- Freebury-Jones, Darren. 2016. "Exploring co-authorship in *2 Henry VI*." *Journal of Early Modern Studies* 5 (March): 201–16. <https://doi.org/10.13128/JEMS-2279-7149-18089>.
- . 2017. "Did Shakespeare really co-write *2 Henry VI* with Marlowe?" *ANQ: A Quarterly Journal of Short Articles, Notes and Reviews* 30 (3): 137–41. <https://doi.org/10.1080/0895769X.2017.1295360>.
- . 2019. "'When a man hath a familiar style': An introduction to authorship studies in early modern drama and literature." *ANQ: A Quarterly Journal of Short Articles, Notes and Reviews* 33 (2–3): 1–10. <https://doi.org/10.1080/0895769X.2019.1677210>.
- . 2021. "Introduction: Computational approaches to examining early modern texts." *Early Modern Digital Review* 4 (4): 191–95. <https://doi.org/10.33137/rr.v4i4.38645>.
- . 2022. *Shakespeare's Tutor: The Influence of Thomas Kyd*. Manchester: Manchester University Press.
- Giddens, Eugene. 2011. *How to Read a Shakespearean Play Text*. Cambridge: Cambridge University Press.
- Gossett, Suzanne. 2022. *Shakespeare and Textual Theory*. London: The Arden Shakespeare.
- Greatley-Hirsch, Brett. 2021. "Computational studies." In *The Arden Research Handbook of Contemporary Shakespeare Criticism*, edited by Evelyn Gajowski, 205–21. London: The Arden Shakespeare.
- Greenblatt, Stephen. 1997. "General introduction." In *The Norton Shakespeare*, edited by Stephen Greenblatt, 1–76. New York: W.W. Norton and Company.
- Hadfield, Andrew. 2021. "Marlowe and Nashe." *English Literary Renaissance* 51 (2): 190–216. <https://doi.org/10.1086/713483>.
- Hirschfeld, Heather. 2015. "Playwriting in Shakespeare's time: Authorship, collaboration, and attribution." In *Shakespeare and Textual Studies*, edited by Margaret Jane Kidnie and Sonia Massai, 13–26. Cambridge: Cambridge University Press.
- Hooks, Adam G. 2016. "The author being dead." In *Shakespeare in Our Time: A Shakespeare Association of America Collection*, edited by Dymphna Callaghan and Suzanne Gossett, 153–57. London and New York: Bloomsbury Arden Shakespeare.

- Hopkins, Lisa. 2005. *A Christopher Marlowe Chronology*. Basingstoke: Palgrave Macmillan.
- Jackson, MacDonald P. 2017. "Shakespeare, *Arden of Faversham*, and *A Lover's Complaint*: A review of reviews." In *The New Oxford Shakespeare: Authorship Companion*, edited by Gary Taylor and Gabriel Egan, 123–35. Oxford: Oxford University Press.
- Jowett, John. 2007. *Shakespeare and Text*. Oxford: Oxford University Press.
- . 2021. "The early printed texts of Shakespeare." In *The Arden Research Handbook of Shakespeare and Textual Studies*, edited by Lukas Erne, 71–93. London: The Arden Shakespeare.
- Knowles, Ronald. 2001. "Introduction." In *King Henry VI: Part 2* by William Shakespeare, edited by Ronald Knowles, 1–141. London: The Arden Shakespeare.
- Love, Harold. 2002. *Attributing Authorship: An Introduction*. Cambridge: Cambridge University Press.
- McInnis, David. 2015. "Marlowe and electronic resources." In *Christopher Marlowe at 450*, edited by Sara Munson Deats and Robert A. Logan, 309–26. Farnham: Ashgate.
- McInnis, David, and Matthew Steggle. 2014. "Introduction: Nothing will come of nothing? Or, what can we learn from plays that don't exist?" In *Lost Plays in Shakespeare's England*, edited by David McInnis and Matthew Steggle, 1–14. Cham: Palgrave Macmillan.
- . 2020. "Introduction." In *Tamburlaine: A Critical Reader*, edited by David McInnis, 1–18. London: The Arden Shakespeare.
- . 2021. *Shakespeare and Lost Plays: Reimagining Drama in Early Modern England*. Cambridge: Cambridge University Press.
- Murphy, Andrew. 2007. "Introduction: What happens in *Hamlet*?" In *A Concise Companion to Shakespeare and the Text*, edited by Andrew Murphy, 1–14. Malden: Blackwell.
- Rudman, Joseph. 2016. "Non-traditional authorship attribution studies of William Shakespeare's canon: Some caveats." *Journal of Early Modern Studies* 5 (March): 307–28. <https://doi.org/10.13128/JEMS-2279-7149-18094>.
- . 2019. "*The New Oxford Shakespeare Authorship Companion*." In *Digital Scholarship in the Humanities* 34 (3): 703–5. <https://doi.org/10.1093/llc/fqz044>.
- Sanders, Norman. 1990. "Shakespeare's Text." In *Shakespeare: A Bibliographical Guide*, New Edition, edited by Stanley Wells, 17–35. Oxford: Clarendon Press.
- Schoenbaum, Samuel. 1966. *Internal Evidence and Elizabethan Dramatic Authorship: An Essay in Literary History and Method*. Evanston: Northwestern University Press.
- Sharpe, Will. 2014. "Framing Shakespeare's collaborative authorship." *Shakespeare Survey* 67: 29–43. <https://doi.org/10.1017/SSO9781107775572.003>.
- Shrank, Cathy L., and Paul Werstine. 2021. "The Shakespeare manuscripts." In *The Arden Research Handbook of Shakespeare and Textual Studies*, edited by Lukas Erne, 53–70. London: The Arden Shakespeare.
- Smith, Emma. 2021. "Shakespeare's early modern books: Printing, paratext and text." In *The Arden Research Handbook of Shakespeare and Textual Studies*, edited by Lukas Erne, 94–110. London: The Arden Shakespeare.
- Stallybrass, Peter, and Roger Chartier. 2007. "Reading and authorship: The circulation of Shakespeare 1590–1619." In *A Concise Companion to Shakespeare and the Text*, edited by Andrew Murphy, 35–56. Malden: Blackwell.
- Stern, Tiffany. 2004. *Making Shakespeare: The Pressures of Stage and Page*. London: Routledge.
- . 2009. *Documents of Performance in Early Modern England*. Cambridge: Cambridge University Press.
- Taylor, Gary. 1987a. "General introduction." In *The Oxford Shakespeare: A Textual Companion*, edited by Stanley Wells, Gary Taylor, John Jowett, and William Montgomery, 1–68. Oxford: Oxford University Press.
- . 1987b. "The canon and chronology of Shakespeare's plays." In *The Oxford Shakespeare: A Textual Companion*, edited by Stanley Wells, Gary Taylor, John Jowett, and William Montgomery, 69–144. Oxford: Oxford University Press.
- . 2016. "Collaboration 2016." In *Shakespeare in Our Time: A Shakespeare Association of America Collection*, edited by Dymphna Callaghan and Suzanne Gossett, 141–49. London: Bloomsbury Arden Shakespeare.

- . 2017. "Artiginality: Authorship after postmodernism." In *The New Oxford Shakespeare: Authorship Companion*, edited by Gary Taylor and Gabriel Egan, 3–26. Oxford: Oxford University Press.
- Taylor, Gary, John Jowett, Terri Bourus, and Gabriel Egan. 2017. "General editors' preface: A complete works." In *New Oxford Shakespeare: Authorship Companion*, edited by Gary Tayler and Gabriel Egan, v–vi. Oxford: Oxford University Press.
- Taylor, Gary, and Rory Loughnane. 2017. "The canon and chronology of Shakespeare's works." In *The New Oxford Shakespeare: Authorship Companion*, edited by Gary Taylor and Gabriel Egan, 417–602. Oxford: Oxford University Press.
- "The first page of Shakespeare's *Henry the Sixth, Part I*. [1623]." 2015. *Wikimedia Commons*. <https://commons.wikimedia.org/wiki/File:First-page-first-folio-1henry6.jpg>.
- "Title page of the first quarto of *Henry VI* [1594]." 2011. *Wikimedia Commons*. [https://commons.wikimedia.org/wiki/File:Henry\\_VI\\_pt\\_2\\_quarto.jpg](https://commons.wikimedia.org/wiki/File:Henry_VI_pt_2_quarto.jpg).
- "Title page of the play now known as *Henry VI, Part 3* [1595]." 2011. *Wikimedia Commons*. [https://commons.wikimedia.org/wiki/File:3H6\\_O\\_TP\\_1595.jpg](https://commons.wikimedia.org/wiki/File:3H6_O_TP_1595.jpg).
- Vickers, Brian. 2008. "Thomas Kyd, secret sharer." *Times Literary Supplement*, April 13, 2008, 13–15.
- Williams, William Proctor. 2018. "Gary Taylor and Gabriel Egan (eds), *The New Oxford Shakespeare: Authorship Companion*." *Notes and Queries* 65 (1): 131–34. <https://doi.org/10.1093/notesj/gjx234>.



## “Slings and Arrows of Outrageous Fortune”: Post-Pandemic Anglophone Theatre and Drama

### ABSTRACT

Since the days of their conception and for most of their history, theatre institutions and the dramatic genre have indelibly reflected their immediate socio-historic contexts, including epidemics. Although forced to close for a full year during the 1918 outbreak of Spanish flu, however, modern Anglophone theatres and authors deliberately avoided exploiting the pandemic in their works. Conversely, the COVID-19 pandemic directly affected the birthing of new genres and individual plays that included it in their settings and contents, and also motivated discussions on the future of dramatic literature and theatre establishments, particularly with regard to hybrid drama. Building on its author's previous research, this paper examines British and American dramatic literature and theatre establishments one year after the end of the pandemic, to detect whether Anglophone drama has embraced the new genres, and whether its authors have continued to reflect the pandemic in their works.

**Keywords:** Anglophone drama, hybrid drama, contemporary drama and theatre, theatre studies

## “Puštic prš proži nezaslišana usoda”: anglofonsko gledališče in dramatika po koncu pandemije

### IZVLEČEK

Gledališča in dramatika sta že od svojega začetka in skozi večji del svoje zgodovine odsevsakokratnega družbeno-zgodovinskega konteksta, vključno z epidemijami. Čeprav so morala takratna gledališča v anglofonskih deželah med izbruhom gripe leta 1918 prvič v novejši zgodovini za celo leto zapreti svoja vrata, so se avtorji v svojih delih namenoma izogibali igranju na karto pandemije. Po drugi strani so med pandemijo COVID-19 nastali novi žanri in dramska dela, v katerih se pandemija pojavlja posredno ali neposredno, obenem pa je sprožila tudi razprave o prihodnosti dramatike in gledališča, zlasti o hibridni dramatiki. Avtorica v svojem prispevku nadaljuje obravnavo britanske in ameriške dramatike ter gledališč leto dni po koncu pandemije ter išče odgovor na vprašanje, ali so se novi žanri uveljavili in ali pandemija v novih delih še vedno igra pomembno vlogo.

**Ključne besede:** anglofonska dramatika, hibridna drama, sodobna dramatika in gledališče, gledališke študije

# 1 Introduction<sup>1</sup>

Theatre institutions and dramatic genres have gone through manifold changes during their centuries-long history, partly as a result of extrinsic socio-historic factors, including wars and various epidemics. In its manifestations both as a performing art (a staged play) and a form of literature (a playtext), the dramatic genre has always reacted to its surroundings, and has to some extent been defined in accordance with socio-historic and geo-political expectations. Within such horizons and contexts the genre has perpetually been recognized, but not always saluted, as a highly relevant means of communal strengthening and social conditioning; a palpable and more captivating form of education; a medium of societal criticism; an efficient instrument of socio-political propaganda; and a method of individual and/or collective consolation. Further, dramas and theatres have always been active promoters of change, especially in moments of deep civilisational crises.

Dramatic authors, regardless of their (sub)genre, have frequently referenced and incorporated local and global historic events and daily geopolitics into their plays. Sometimes they do so literally and in passing, and at other times more substantially, turning such events and details into metaphors and symbols of existential and metaphysical dilemmas, and prompting an internal discussion, or even an actual (re)action from audiences (and lawmakers) on a particular issue. Ancient classical tragedians habitually set the action of their plays against the backdrop of (real) wars and epidemics. *Oedipus Rex* (c. 430 BCE), for example, one of the best known Ancient Greek tragedies, sets its quest for the perpetrator of a regicide against the backdrop of a pestilence, possibly invoking the plague of Athens (Kousoulis et al. 2012). These events of considerable death and destruction would have been understood as symptoms of moral deprivation and civilisational decay that could offer a valuable lesson to spectators. Further, they would have symbolically represented critical times of (self-)scrutiny, in which a different, more functional (if not better) society and social order could be conceived. In other words: fictionalised re-presentations of events of acute crises, whether wars or epidemics, would have been employed by authors and understood by audiences as “moment[s] of reckoning, the space when the intervention might still take place” (Angelaki 2017, 2).

In a similar manner, Anglophone drama (especially since the rise of secular Tudor drama and the works of its bard, William Shakespeare), has always reacted to and referenced its immediate socio-historic contexts. It has used wars, natural disasters and, sporadically, epidemics, as pretexts for promoting or probing current ideologies, myths and politics,<sup>2</sup> either to uphold the status quo or obliquely offer social criticism. Born in the midst of the bubonic plague, and subject to the consequences of constant outbreaks of the disease – which impeded life in Elizabethan England, including its theatrical activities – Shakespeare used the epidemic as a pretext for a play’s action only once, although he referred to pestilence frequently in other texts (Tichenor 2021; Young and Hagan 2020). The one occasion on which Shakespeare wove the plague into a plot was when it prevented Friar John from delivering the letter to Romeo (*Romeo and Juliet* 1982, 5.2.). In all his other plays, the plague (in whatever variant)

---

<sup>1</sup> An early version of the paper was presented at the 6th International SDAŠ Conference in September 2023.

<sup>2</sup> Literary texts (of any genre) have always been (mis)used for (geo)political purposes and propaganda. A studious and inspiring text on the topic can be found in Hazemali and Onič (2023).

is given a symbolic meaning, either as a symptom of personal or societal corruption, or a threat of impending catastrophe caused by the depravity of its protagonists (who are seen and depicted as social agents). Within the latter context, these events also represented an opportunity for social intervention, and could eventually be interpreted as “agents of change and transformation” (Fernández-Caparrós and Brígido-Corachán 2017, xvii).

The present paper is a continuation of its author’s research into whether and how contemporary American theatres and dramatic literature have reacted to the recent (COVID-19) pandemic, and is part of an ongoing study. The previous research was limited by its relatively short time span and may have led to inaccurate observations and initial conclusions. Therefore, this current research continues to observe and examine American and British dramatic literature and theatre establishments one year after the pandemic, with the primary aim of detecting whether Anglophone drama has embraced the new hybrid genres prompted by the health mandates in place during the COVID-19 pandemic. Its secondary intention is to inspect whether authors and directors have continued to reference the pandemic in their work, and to what degree. Its third, and most important, intention is to establish whether the crisis brought about any lasting transformations and interventions in Anglophone theatre and drama.

## 2 Previous Research

Theatres were among the first social institutions to be locked down at the outbreak of the COVID-19 pandemic, and subsequently experienced the most extended complete shut-down in their history, including the period at the beginning of the 20th century when the 1918 influenza pandemic halted the performing arts for a year (Canning 2020; Colleary 2021). Unlike other educational and cultural organizations, such as schools, universities, museums, and even the film industry, theatres and the dramatic genre are inherently dependent on the presence and proximity of audiences to be functional and viable. The health protocols mandated during the COVID-19 pandemic were thus in direct opposition to the intrinsic essence of the performing arts and staged drama, and an impediment to their existence. Additionally, by the 2020s most contemporary Anglophone theatres had completed a developmental cycle and arrived at the needs-assessment phase. Having previously received long-called-for demands for better social representation, inclusion and greater equity for the marginalised – among audiences as well as among dramatists, troupes and various theatre professions – Anglophone theatres were being compelled to re-evaluate their respective agendas, organisational structures, and long-term sustainability. The author’s aforementioned paper, published in late 2022, examined the condition of US theatres and drama in the contemporary era of crises, observed their immediate reactions to the global pandemic, and noted the ensuing changes. The premise of the research was that crises are “agents of change and transformation” (Fernández-Caparrós and Brígido-Corachán 2017, xvii), and that US theatre and drama could use the COVID-19 pandemic as a “space [of] intervention” (Angelaki 2017, 2), in order to tackle and successfully deal with its own crisis of representation and parity.

The early findings were that theatre companies around the world employed accessible digital and streaming platforms and/or directed their attention to the older, relatively overlooked, genre of the radio-drama (Lunden 2020a; Lunden 2020b; Langston 2022), so

that: “Productions did not stop completely, but instead went online” (Langston 2022, n.p.). This initial reaction from performers, authors and theatre professionals to work remotely from the safety of their homes and to experiment with contemporary technologies led to the burgeoning of new hybrid forms of theatre and drama. Because of the imposed health mandates, the creative survival strategies of those employed in theatre, and the deliverance the technologies provided, the dramatic genre has expanded since March 2020 to include sub-genres such as digital play-readings, micro-plays, audio theatres, Zoom plays, Instagram monologues, YouTube shorts and other full-scale digital(ised) (quasi)live-theatre subgenres. Major weaknesses of such developments, however, include the “lack of intimacy and dynamism of online theatre” (Langston 2022, n.p.), and eventual “Zoom fatigue” (Blake 2020, n.p.) or “digital theatre exhaustion”, a feeling that arose with the realisation that returning to physical theatres and live shows would take much longer than imagined.

The second observation from the earlier research was that digital performances and streamed and live plays in 2020 and the first half of 2021 (i.e., in the 2019–2020 and 2020–2021 seasons) only sporadically exploited the pandemic in their plots and source materials. This was the case in Jessica Blank and Erik Jensen’s *The Line* (a YouTube-streamed docu-drama that portrayed frontline medical workers battling the pandemic); Richard Nelson’s *What We Need to Talk About* (a Zoom play, and the sequel to Nelson’s *The Apple Family* plays, which also addressed life during the COVID-19 pandemic); and Kristina Wong’s *Sweatshop Overlord* (an Asian-American comic take on the COVID-19 experience). In most cases, however, theatre in 2020–2021 preferred to alleviate the stress of living with the pandemic and lockdown and offer a form of escapism through musical drama, comic satires, and revivals of national and world classics whose thematic frameworks steered clear of the pandemic. Still, *The Great Work Begins* (a streamed benefit play-reading of scenes from Tony Kushner’s *Angels in America*, presented by amFAR and directed by Ellie Kanner) points to the early apprehension that the health crisis would hit historically marginalised individuals and communities (LGBTQ+, BIPOC, women, and the poor) disproportionately hard, forcing theatre professionals to take prompt action.

Another key finding of the research was that the COVID-19 pandemic propelled theatres to begin a much-needed reckoning process. This was driven in part by the previously mentioned demands for a reappraisal of their programming and structures, and other social and historic events, such as the 2020 Black Lives Matter riots (sparked by a spate of excessive police violence against Black Americans), racially-motivated acts of aggression and violence against Asian-Americans, and reports and data that supported the further social marginalization of BIPOC, women and LGBTQ+ individuals, who were more likely to be furloughed or laid off, and had limited (if any) access to medical assistance. Discussions on how to achieve better societal representation, inclusion and equity brought a series of early propositions, the details of which can be found in the previous research (Čirić-Fazlija 2022). A comparative assessment of the dramatic genres, emergent authors and modes of societal representation that re-directed the trajectory of US theatre in the 20th and 21st centuries, however, indicates that a greater transformation of US drama and theatre took place in earlier decades rather than more recently. This is despite the rise of new/hybrid genres necessitated by the recent pandemic, and the increased number of plays by American women staged in commercial and independent theatres.

### 3 Research Methodology

To conduct the current research, the author accessed and examined a variety of sources that focus on theatres and performances in 2022 in the US and UK, in independent non-profit and commercial institutions. The sources fell into two main categories. The first comprised reports and surveys from national theatre associations/organisations, including the Theatre Communication Group's reports and editorials (Fonner et al. 2022; Weinert-Kendt 2023); the LA Performing Arts Survey (Star Insights 2022); and writings by individual reporters and theatre columnists (Paulson and Hernández 2022; Gelt 2022; Veltman 2023; Mason 2023; Farber 2023), all of whom cited and relied on specialised surveys and analytical data on theatres.

The second category of resources consisted of “best of” theatre lists, compiled by established theatre critics and columnists such as Jackson McHenry (2022) of the *Vulture*, Jesse Green (Green et al. 2022) and Maya Philips (2022) of *The New York Times*, Peter Marks (2022b) of *The Washington Post*, Adam Feldman (2022) of *Time Out New York*, David Gordon (2022) of *Theatermania*, Charles McNulty of *LA Times* (2022b), Holly Williams, Jessie Thompson and Isobel Lewis (2022) of *The Independent*, Arifa Akbar (2022e) of *The Guardian*, Dominic Cavendish (2022) of *The Telegraph*, Susana Clapp (2022) of *The Observer*, Sam Marlowe (2022) of *The Stage*, Andrzej Lukowski (2022) of *Time Out*, Sarah Hemming (2022) of *The Financial Times*, and Daisy Watford of *Matadornetwork* (2022). At times various theatre websites were accessed for the purpose of clarification, but their information on the season's programmes was not taken into account, as it often proved changeable and incomplete.

### 4 Research Findings: Reports and Surveys

Most of the aforementioned reports and surveys testified to the fact that theatres have suffered significant losses, not only in financial terms but more significantly with regard to their personnel, as “unemployment in the arts doubled the nation average during the pandemic, spiking at 30%”, especially among BIPOC workers (Gelt 2022, n.p.; SMU DataArts 2023). Moreover, the frequent suspensions and cancellations of the performances that occurred in 2021 (either when a new variant of COVID-19 emerged, or companies' staff fell ill), the generational flux of the audience, and changing social habits have meant that the return to live, in-person theatre has been surprisingly slow and faltering. Additionally, by the end of 2022 federal grants and relief funds in the US, such as the Payroll Protection Program (PPP), Shuttered Venues Operating Grant (SVOG), and funding from the Los Angeles County Department of Arts and Culture (LACDAC), if successfully won, dried out, had mostly been spent on “operating and reopening efforts”, and to “cover losses incurred due to pandemic-related shutdowns” (Star Insights 2022, 9). The funds generally did not prevent institutions from losing employees, or from having to discontinue their activities. Further, although individual (and last-minute) ticket sales did not drop, subscriptions fell significantly (30–60%), and the general political and economic volatility combined with high inflation have reduced funding from benefactors and corporations. These factors combined have eventually led to fewer performances being given, shorter runs, further losses of performing arts jobs, and the permanent closure of some (non-profit) theatres. This has had far-reaching implications

on theatre programming, as companies have had to prioritise, placing sustainability over try-out/new drama, as explained by Snehal Desai:

I think what we are seeing is our audiences are coming back, but they are being really selective, which makes it hard to stage riskier plays or challenging material – or to do new work or introduce new artists. (Gelt 2022, n.p.)

The realisation that it will take a much longer time for the performing arts to recover from the pandemic (another three to five years, according to the artistic and theatre executives interviewed), and the generational shift among audiences<sup>3</sup> (which necessarily implies different habits, interests and cultural tastes) do, however, have some upsides. First of all, theatres have started to restructure and re-think their organisational policies and artistic agendas, in order to devise new strategies and methods of attracting audiences and adapting to the newest normal in the “era of diminished expectations” (Weinert-Kendt 2023, n.p.). Some have dismissed their boards and traditional hierarchies, others have changed their marketing strategies, and others still have diversified into programming “outside traditional, proscenium-based theatre” (Weinert-Kendt 2023, n.p.). The development of programmes beyond the bounds of traditional theatre implies the inclusion of immersive, virtual and augmented theatre, which, according to Alan Brown of the WulfBrown consultancy, “the public has embraced [...] and the commercial producers are running away with millions of dollars in demand for them”, even if a certain number of artistic directors still prefer “to do important theatrical work on their mainstages [sic] for an audience of critics” (quoted in Weinert-Kendt 2023, n.p.). The second upside is that theatres have undertaken the necessary reckoning and have already intervened in their programmes, making them increasingly diverse and inclusive, with more texts and performances by the traditionally marginalised (women, BIPOC, LGBTQ+). This is already apparent in the “best of” lists, which have correspondingly become more inclusive.

## 5 Findings from the “Best of” Lists

The “best of” lists compiled by the previously listed theatre reviewers, critics and columnists reveal that theatres and audiences in 2022 preferred revivals, Broadway transfers and new productions of older national dramas and world classics. As expected, musical dramas<sup>4</sup> such as *Funny Girl*,<sup>5</sup> *A Strange Loop*,<sup>6</sup> *Kimberly Akimbo*,<sup>7</sup> *Into the Woods*, *Intimate Apparel*,<sup>8</sup> *Cyrano*

---

<sup>3</sup> This phenomenon is partly natural and partly the result of the pandemic, with older generations observing safety precautions and health mandates more strictly.

<sup>4</sup> More information on songs as generic elements in film musicals can be found in Plemenitaš (2016).

<sup>5</sup> This 2022 performance, starring Lea Michele, was the first Broadway revival of the 1964 musical by Isobel Lennart (book author), Bob Merrill (lyrics) and Jule Styne (score). It presents the career and personal life of comedienne Fanny Brice of *Ziegfeld Follies*.

<sup>6</sup> Michael R. Jackson's meta-theatrical musical dramatises the daily and social struggles and concerns of a gay Black author writing a musical about a gay Black author writing a musical. First premiered in 2019 off-Broadway, the piece won the 2020 Pulitzer Prize for Drama (for more information, see: Pulitzer Prizes 2020), and Best Book of a Musical and Best Musical at the 75th Tony Awards in 2022, after being nominated in 11 categories. The 2022 performance was its first Broadway premiere. For more information, see: Wiegand (2022) and the “Winners/2022” section on the Tony Awards website (Broadway League 2022).

<sup>7</sup> For more information on the play and the 2022 performance, see: Ćirić-Fazlija (2022, 67), and Kennedy (2022).

<sup>8</sup> *Intimate Apparel* is a 2004 playtext by Lynn Nottage, which foregrounds the intersectionality of race, gender and class in its representation of an African-American working-class woman at the turn of the 20th century. It was not

de Bergerac, *The Life*,<sup>9</sup> *The Bed Wetter*,<sup>10</sup> *Guys and Dolls*, *AD16*,<sup>11</sup> *American Prophet*,<sup>12</sup> *Merrily We Roll Along*, *♫Juliet*, *Some Like It Hot*, *MJ: The Musical*, *Back to the Future*, *The Hang*, and the unavoidable *Oklahoma!* dominated, followed by dark comedies, (comic) satires, parodies<sup>13</sup> and new stagings of Shakespeare,<sup>14</sup> and quite a few docudramas and mono-dramas/solos.

If the authorship and thematic framework of the plays are considered, it must be concluded that a palpable increase in racial, gender and cultural diversity and parity is evident. This is visible in the number of playtexts and performances by or about BIPOC, women, and LGBTQ+ individuals. The scripted dramas include: Alice Childress's *Trouble in Mind* and *Wedding Band* (portraying the difficulties of the interracial love and marriage of a black woman and a white man in the American South in 1918, told from the female protagonist's point of view);<sup>15</sup> Lynn Nottage's *Intimate Apparel*; Susan-Lori Parks' *Topdog/Underdog* (highly original in its treatment of drama and stage conventions, and postmodernist in its ironic re-visitation of American history, the play premiered in 2001 off-Broadway, but was transferred to Broadway the following year);<sup>16</sup> Pearl Cleage's *Blues for the Alabama Sky* (this 1995 play, set in 1930s Harlem, reveals its author's complex probing of US historiographical narratives, and her attempt to recover African-American women from the neglect and oblivion of official history);<sup>17</sup> Adrienne Kennedy's *Ohio State Murders* (1991/1992 play, which thematises a race-related crime, had its Broadway premiere in 2022);<sup>18</sup> Paula Vogel's *How I Learned to Drive* (dramatising issues of sexual abuse and paedophilia within the larger coming-of-age narrative

---

originally devised as a musical play, although the performance mentioned here was staged as one: "*Intimate Apparel* began its life as a popular play, but it was the brilliant composer Ricky Ian Gordon who invited me to consider adapting it into an opera. He saw something epic and expansive in the life of Esther that he felt demanded to be sung, and with his loving guidance, I was able to write my first libretto" (Nottage 2022, n.p.).

<sup>9</sup> A revival of David Newman, Ira Gasman and Cy Coleman's 1990 classic, which depicts the street life of colourful characters of sex workers in 1980s New York. It had its first Broadway premiere in 1997, and won two of the twelve Tony awards it was nominated for that year (see Concord Theatricals 2023; and Playbill 2023).

<sup>10</sup> A new musical by Sarah Silverman, Joshua Harmon and Adam Schlesinger, which was set to premiere in 2020 and is based on Silverman's 2010 memoir. For more information, see: Jessie Green's review (Green 2022b), and the Atlantic Theatre Company's website (Atlantic Theatre 2022).

<sup>11</sup> Cinco Paul and Bekah Brunstetter's new musical (which was both streamed online and performed in-person in early 2022) gives a humorous take on the biblical story of Mary Magdalene and Jesus as teenagers. For trailers and more extensive information, see Marks (2022a) and Hall (2022).

<sup>12</sup> A new musical by Charles Randolph-Wright and Marcus Hummon, which presents the life and activism of Frederick Douglass. For additional information, see: the American Prophet website (ShowTown Theatricals 2022), and David Smith's piece (2022).

<sup>13</sup> For example: *Fat Ham*, and the aforementioned *♫Juliet*.

<sup>14</sup> Shakespeare's *Othello*, *Hamlet*, *Henry V*, *Merchant of Venice*, and *Much Ado About Nothing* particularly captured the attention of directors and audiences in the season(s) examined.

<sup>15</sup> *The Wedding Band* was written in 1962, preceding anti-miscegenation laws by five years. The premiere listed here was the play's "first major New York revival in 50 years" (Green 2022a).

<sup>16</sup> *Topdog/Underdog* was awarded a Pulitzer Prize in Drama in 2002 and celebrated as "A darkly comic fable of brotherly love and family identity, [...] tell[ing] the story of Lincoln and Booth, two brothers whose names, given to them as a joke, foretell a lifetime of sibling rivalry and resentment" (Pulitzer Prizes 2002, n.p.).

<sup>17</sup> As the author states: "I am writing to expose and explore the point where racism and sexism meet. I am writing to help myself understand the full effects of being black and female in a culture that is both racist and sexist" (Cleage quoted in Anderson 2008, 17).

<sup>18</sup> As the unnamed author of the promotional material explains: "This haunting, provocative mystery – the Broadway debut of legendary playwright Adrienne Kennedy – concerns a famous writer who returns to her alma mater to finally reveal the truth of what happened when she was a student there" (Concord Theatricals 2022, n.p.).

of a naïve female character, in a poignant and harrowing dissection of its topics);<sup>19</sup> Caryl Churchill's *A Number* (2002 play that probes the issue of cloning and the potential to atone for past mistakes through scientific experiment);<sup>20</sup> Helen Edmundson's *A Small Island* (two-act scripted drama which portrays the lives of three Windrush generation women migrants from the Caribbean to the UK);<sup>21</sup> Aya Ogawa's *Nosebleed* (an autobiographical piece by the playwright-cum-director presented as a work-in-progress in 2019, and premiered in autumn 2021);<sup>22</sup> Suzie Miller's highly acclaimed *Prima Facie* (a one-act, solo drama which presents a female lawyer who specialises in defending men from sexual abuse charges, at the tragic moment of having been sexually attacked herself);<sup>23</sup> August Wilson's *The Piano Lesson*;<sup>24</sup> J.E. Cooper's *Ain't No Mo'* (the 2019 play of the Black author-cum-performer that satirically examines racialised American society);<sup>25</sup> R. Calais Cameron's *For Black Boys Who Have Considered Suicide When The Hue Gets Too Heavy* (highly lauded play which problematises the (self)perception of Black masculinity);<sup>26</sup> Michael R. Jackson's *A Strange Loop*; Richard Greenberg's *Take Me Out* (discussing class-, race- and gender-based chauvinism in the sports industry);<sup>27</sup> Mike Bartlett's *Cock*;<sup>28</sup> Jeremy O. Harris's *Daddy: A Melodrama*;<sup>29</sup> Katie Elin-Salt's *Celebrated Virgins*;<sup>30</sup> Mathew Lopez's *The Inheritance*;<sup>31</sup> and Sami Ibrahim's *Two Palestinians Go Dogging*. The latter, although grounded in the 2022 reality of the cold-blooded murder of an Al Jazeera journalist, was sadly to turn prophetic by the end of 2023.<sup>32</sup> Among the

<sup>19</sup> A highly relevant and acclaimed work by an American woman dramatist, Vogel's play premiered in 1997, and received a Pulitzer Prize for Drama in 1998 (Pulitzer Prize 1998).

<sup>20</sup> Churchill is an established and acclaimed British female dramatist. This play was revived numerous times, the last in 2022 at the Old Vic in London (The Old Vic 2022).

<sup>21</sup> First produced in 2019, Edmundson's play is based on Andrea Levy's 2004 novel. For more information on the play and its context, see: Sierz (2019) and BBC News (2023) respectively.

<sup>22</sup> *Nosebleed* dramatises the "insurmountable cultural and generational gap between Aya and their father" (Ogawa n.d., n.p.).

<sup>23</sup> This was the West End debut of Suzie Miller's drama, which had its Broadway premiere in 2023. As Arifa Akbar (2022b, n.p.) states: "*Prima Facie's* final messages are urgent in highlighting who our laws fail to protect. If they are delivered in hammer blows, there is power in hearing them spoken on a West End stage..."

<sup>24</sup> A revival of the 1987 play by highly revered contemporary African American dramatists, set in 1936 Pittsburgh.

<sup>25</sup> The play's short run on Broadway testifies less to the quality of the piece and more to the difficulties US theatres had when trying to introduce new works in the early post-pandemic period. For more information, see: Huston (2022) and Oladipo (2022).

<sup>26</sup> Inspired by Ntozake Shange's 1975 choreopoem, the play premiered in October 2021 and was transferred to the Royal Court Theatre in 2022. It has had a West End theatre run in 2023 (Hall 2023; Ryan 2023).

<sup>27</sup> A Broadway revival of the 2002 play. For information on the 2022 revival, which counts the Tony for Best Play Revival among its many awards, see: Towers (2022) and Frank (2022).

<sup>28</sup> The play premiered at the Royal Court Theatre in 2009, and portrays the dilemma of a gay man who "accidentally fall[s] in love with a woman" (Dramatists Play Service 2009, n.p.).

<sup>29</sup> A play that foregrounds the issues of race, gender and sexuality embedded in the politics of modern patronage, *Daddy: A Melodrama* was first performed off-Broadway in 2019 (Lewis 2019; Akbar 2022a). For information on its 2022 UK premiere, see: Akbar (2022a) and "Daddy: A Melodrama" (Almeida Theatre 2022).

<sup>30</sup> A new play by the actress and emerging author from South Wales, *Celebrated Virgins* presents a fictionalised history of 'The Ladies of Llangollen', two women from the upper echelons of Irish society, shunned by early 19th-century society because of their lesbian relationship (see Gareth Williams's interview with the dramatist [Williams 2022]). The play "attempts to reclaim their story on their own terms" (Llyr Evans 2022, n.p.).

<sup>31</sup> This revival of the 2019 LGBTQ+ drama premiered on Broadway. It was inspired by Tony Kushner's *Angels in America* and E.M. Forster's *Howards End*. For additional information on the original production and most recent revival, see: Travers (2019) and McNulty (2022a), respectively.

<sup>32</sup> For more information on the play and its coincidence with the killing of Shireen Abu Aqleh, see Arifa Akbar's review (2022c); for information on the murder of the famous Al Jazeera journalist see Bethan McKernan and Sufian Taha's piece (2023) in *The Guardian*.

plays by BIPOC, women, and/or LGBTQ+ authors (and which are in part intersectional) two are particularly interesting: *The Father and the Assassin*, a new play by the Indian woman dramatist Anupama Chandrasekhar (the first International-Playwright-in-Residence at the National Theatre in London), which unravels the history of British colonialism through its focalisation on the killing of Ghandi, and is “a story of division and whipped-up animosities that has its roots in colonialism and is repeating itself throughout the world today” (Armitstead 2022, n.p.); and *The P Word* by British actor-cum-author of Asian/Middle-Eastern origin Waleed Akhtar (Akhtar 2022), “a consciousness-raising play” (Akbar 2022d, n.p.) that foregrounds gender-based prosecution, homophobic castigation, racial chauvinism inside the LGBTQ+ community, and global migration in the context of the British asylum-seeking system.<sup>33</sup> Curiously, among the more anticipated revived classics (Shakespeare, Shaw, Miller and Williams), are a few examples of theatre of the absurd, and meta-theatrical and experimental drama, such as Ionesco’s *The Chairs*, Albee’s *Three Tall Women* and Wilder’s *The Skin of Our Teeth*. These plays are all strongly existentialist and metaphysical in their contents and thematic frameworks, yet they (especially the latter) remain optimistic in tone. It appears that the intimate and collective experience of living with (and surviving) a global health crisis and its (at times inconsistent) mandates have had an impact on programme selectors, who reached for the older plays audiences and artists could learn from and find comfort in.

Quite a number of new plays and premieres caught the attention of critics and reviewers, including: *A Case for the Existence of God*, a play on the issue of custody and single fathers’ battles for their children; *Camp Siegfried*, a piece which foregrounds camps for young German-Americans that indoctrinated children with Nazi ideology; *English*, a text that debates global migration crisis, and the Iranian migratory experience and cultural identity; *At the Wedding*, a drama that discusses the LGBTQ+ experience; *The Hang*, an LGBTQ+ performance that combines the musical and jazz cabaret genres; *Here Are the Blueberries*, a docu-drama on Nazi concentration camps; *That Is Not Who I Am*, a mystery-piece on online conspiracy theories; *House of Ife*, which presents a British-Ethiopian family after the death of its eldest son; *Two Palestinians Go Dogging*, which references the Palestinian-Israeli conflict; *Our Generation*, a piece of verbatim theatre<sup>34</sup> on Generation Z’s coming-of-age; *Beautiful Evil Things*, a solo that retells the ancient tragedies of Medusa, the Amazon Queen Penthesilea, seer Cassandra, and Leda, mother of Clytemnestra; and *The Glow*, a sci-fi and folk-myth drama about a woman with supernatural powers. This list compels us to agree with Arifa Akbar that it was “a superb year for the stage” (2022e, n.p.), and with Charles McNulty, who describes the 2021–2022 season as one marked by “theatrical richness”, although it was “a bruiser, economically and emotionally” (2022b, n.p.).

Of all the plays and performances included in these “best of” lists, only three directly referenced the pandemic: a pared-down and modernised staging of Tennessee Williams’s *A Long Day’s Journey into Night*, performed at the Minetta Lane Theatre, directed by Robert O’Hara and

<sup>33</sup> For more information on public reactions to the 2016 migration crises in UK and Slovenia, and on how to use songs to successfully battle xenophobia and teach inclusivity and interculturalism, see Hempkin (2016).

<sup>34</sup> For a definition and explanation of verbatim theatre, see: Geoff Willcocks, “Europe in Flux: Exploring Revolution and Migration in British Plays of the 1990”, in Holdsworth and Luckhurst (2008, 7–25); and Agnes Wolley’s *Contemporary Asylum Narratives: Representing Refugees in the 21st Century* (2014, 117–40).

starring Bill Camp and Elizabeth Marvel (Shaw 2022); Max Webster's modern-dress version of William Shakespeare's *Henry V* for the Donmar Warehouse (and streamed on National Theatre Live), starring Kit Harington (Watford 2022); and Alecky Blythe's *Our Generation*, co-produced by the National Theatre and Chichester Festival Theatre, and directed by Daniel Evans. The first uses the pandemic as a backdrop for its plot; the second references the pandemic briefly in its stage design; and the third, an example of coming-of-age verbatim theatre, weaves the health crisis into its plot, as it dramatises the dilemmas and issues of a dozen Gen Z youths from across the UK over a five-year period, which ended in 2020. Since the first two of the aforementioned texts and performances do not include the COVID-19 pandemic in a substantial way, the next subsection of the paper focuses solely on Alecky Blythe's *Our Generation*.

## 6 *Our Generation: Coming of Age in the Contemporary Era of Crises, the Pandemic Included*

Alecky Blythe's most recent<sup>35</sup> piece in the genre of verbatim theatre, *Our Generation*, came to fruition over a period of five years, starting in 2015, when its author came up with an idea to devise a piece that would reflect the experiences of a "generation of young people in the UK coming of age in the 2020s" (Blythe 2020, "Introduction"). When the concept was first discussed by Blythe and Rufus Norris, the newly-appointed Artistic Director of the National Theatre, neither the author nor her interlocutor could foresee what would be happening by the end date, and the dramatist expressed her concern that such a project would present structural difficulties, because it "could be potentially so freewheeling [...] without a central event to aid knitting it together" (Blythe 2022, "Introduction"). Despite the fact that Norris tried to allay Blythe's fears by citing the unpredictability of the future, suggesting that the intervening years might bring "wars, terrorist attacks, a pandemic" (as quoted in Blythe 2022, "Introduction"), neither could foresee the (specific impact of the) COVID-19 pandemic, and the striking effect it would have on their research subjects. Yet, when the health mandates were imposed in March 2020, the project of tracking and interviewing twelve British youths from around the UK became not only an accurate representation of the authentic predicaments, anxieties, interests, dreams and struggles of a cohort of maturing teenagers, but also a testimony to the strength, adaptability, buoyancy and overall resilience of a generation affected and even disrupted by various manifestations of crises, from substance abuse and mental-health epidemics, to the infodemic, Brexit, and eventually the COVID-19 pandemic.

Blythe selected her subjects carefully, paying attention not only to the age criterion (all were supposed to be in their early teens at the beginning of the project – the youngest was 12, and the oldest 15; see Blythe 2022, "Characters"), but also the criteria of gender, class, race, and

---

<sup>35</sup> An actress, author, screenwriter and director, Blythe began to experiment with verbatim theatre in 2003, when she established the Recorded Delivery theatre company. Since then, she has composed 16 pieces for stage, TV and film, including *Come Out Eli* (2003), *Strawberry Fields* (2004), *Cruising* (2006), *The Girlfriend Experience* (2008), *Do We Look Like Refugees?* (2010), *London Road* (2011), *The Riots: In Their Own Words* (2012) and *Little Revolution* (2014) (United Agents n.d.; Drama Online Library 2023). Her specific process of creating a sub-genre of verbatim theatre, known as "recorded delivery" (in which the actors on stage are directly fed the edited lines of authentic recorded dialogue via an earpiece) has been explained and elaborated on in various articles and interviews. For more information, see: Sarah Lambie (2022), Simon Stephens (n.d.), Natalie Woolman (2012), Philip Fisher (2006), and Alecky Blythe (2022, "Introduction").

nationality, thereby “representing the geographical breadth of the country” (“Introduction”). Although the play lists 94 characters, including the 12 chief protagonists’ familial, scholarly, and/or intimate connections and associations, its focus remains on: Ierum, a 12-year-old North-African British girl; Luan, a 15-year-old British boy of Kosovan descent (two characters from South London); Anabella, a 13-year-old girl of dual heritage; Callum, a 13-year-old white boy (two characters from Belfast); Emily, a 14-year-old white girl; Lucas, a 14-year-old white boy (two characters from Cambridgeshire); Zak, a 15-year-old boy of dual heritage, with an Indian mother; Robyn, a white 15-year-old girl (two characters from Glasgow); Mia, a white 14-year-old girl; Taylor, a white 14-year-old boy with a disability (two characters from North Wales); and Ali and Ayesha, a Birmingham-based Pakistani British brother and sister, aged 15 and 12, respectively. The characters come from various kinds of households, including single-parent and broken homes, and experience myriad challenges. The trials readers witness include age-related worries about school results and GCSE exams; identity issues and the need to fit in; first romantic relationships; rebellion against authority figures; and heart-rending distress and ordeals to which no young adult should be exposed, such as child neglect, substance abuse, unwanted teen pregnancy, stress-induced mental breakdowns, and even physical abuse at the hands of a romantic partner. Although many fragments in Act Four might induce profound sympathy in the reader, among the most distressing series of fragments in this context are those relating to the horrendous experiences of Mia, shown in the following excerpt:

I’ve just had a really, really shit time yeah. (*Pause.*) Got into like really bad, like really really badly abusive relationship for thirteen and a half months. (*Beat.*) Do you remember when you came to see me last when we’re in college? It’s the same lad. He was on bail. So he’s admitted, they [sic] he’s admitted to one allegation which was throwing me round. They’ve got ten charges, then they’ve also now got charges for coercin’ an’ controlling behaviour. He’d ripped all my hair out, left big dents on my knees where he’d like jump on top of me and stuff, kicked me round the plot, smacked my head against the side, this is just in like one day yeah. (*Pause.*) I overdosed yeah. ... (Blythe 2022, 2.4.24.)

The playtext is organized into three parts, and its dramatic dialogue is spread over six acts and 208 episodic scenes. The first two acts are presented in Part One, each of which comprises 28 scenes; the next two acts make up Part Two, and are of different lengths – Act Three comprises 41 scenes; and Act Four 35. Act Five (Part Three) is the longest, comprising 59 scenes, while the last act, Act Six (also in Part Three) is the shortest, with only 17. Most scenes are reasonably short monologues and/or duologues by various of the 12 teens, and appear as juxtaposed fragments, making the reader privy to a specific moment in the life of a character. All scenes are numbered (by Part, Act and Scene) and almost all have a title, which attaches a key descriptive phrase for the experience that dominates the scene and/or reverberates in the episode. The opening fragments in Act One of Part One serve as introductions to the *dramatis personae*, whom the reader follows in their progression and maturation from early teens to young adulthood over the course of the play. A certain number of episodes are devised as ensemble scenes, which place the 12 characters next to each other, offering a dissonance in the individual voices of the generational experience. Such is the last scene in Act Four,

entitled “The Future Ensemble”, in which the 12 characters discuss their individual “hopes and aspirations” for the future (Blythe 2023, 2.4.35.). The episode ends with a ripping sound, signalling an approaching catastrophe, which will put their dreams and plans on hold:

ANNABELLA. I don't like to think about the future, like ever. I literally haven't even thought about what I'm gonna eat for dinner. I just think like, like I'm happy now, you know I'm fifteen like I'm so young, I don't want to think about the future, I want to be hap- happy like now. There's going to be so many surprises in life, there's jus' no point in thinking about it. I wanna get this year over at school and then (*Beat.*) we'll see what happens then.

ANNABELLA *pulls away the gauze as a vicious sound rips through the auditorium. The youngsters are revealed all stood in a line wearing face masks, staring blankly at the audience.* (Blythe 2022, 2.4.35.)

The dramatic tension rises throughout the episode, and the irony is tangible as the scene's closing *didascalie* announces the beginning of the COVID-19 pandemic as one of the “many surprises” Anabelle mentions, without fully grasping her own prophetic words.

The play's longest act, Act Five, is situated in Part Three, and depicts a year in the life of the youngsters, one suspended by the pandemic. It is a vivid reminder of the minutest details of the *annus horribilis* of 2020. The stage direction is decisive about the setting: all characters (and their parents, friends and partners) now appear as if at an extended Zoom conference. They talk about safety measures, health protocols, numbers and percentages of people falling ill and dying, social distancing, the cancelling of their plans and exams, and the difficulty of maintaining relationships while living the “hap-happ[iest]” days of their lives remotely, and in seclusion. Their openness about their individual struggles to maintain a meaningful life is heartfelt yet agonising, and reveals both the fragility and unexpected resilience of the characters at such a precarious (historic) moment, which made them mature overnight. Instead of celebrating momentous birthdays and having graduation parties, starting college and university studies, partaking in the Paralympics (and maybe winning a medal), getting their first “serious” job (one that could help to procure a home), or simply taking a year-long tour of Europe, their rite of passage comes in the form of the COVID-19 pandemic. Alecky Blythe, however, ensures that the audience feels the hope of the generation that was disrupted “when suddenly the world stopped” (“Introduction”). Act Six reveals that all 12 characters have survived the health scare, and come out of it “stronger than [they] thought” (3.6.12.), although they feel they have “aged, more than [they] did, [...], in the previous eighteen ...” (3.6.13.). *Our Generation* is therefore among the rare Anglophone dramatic texts that delves into the representation of the COVID-19 pandemic, and is a lasting testimony to the poignant resilience of humankind in the face of yet another calamity.

## 7 Conclusion

This paper is a part of longitudinal study on Anglophone theatre and its responses to the contemporary era of crises, in particular the health crisis of 2019–2022. Its preliminary results should be read with caution, as the research is not yet complete, and the author intends to continue to observe developments in the American and British dramatic literature and

performing arts scenes in the aftermath of the COVID-19 pandemic, in order to establish its longer-term effects on Anglophone drama and theatre. Still, the preliminary findings do demonstrate that, although not necessarily detectable in the first year of the pandemic (which coincided with the latter half of the 2019–2020 and first half of the 2020–2021 seasons), Anglophone theatres have in the meanwhile taken the crisis as a “moment of reckoning”. In the course of the 2021–2022 season, theatre institutions have intervened significantly in structural and programming decisions, so as to battle exclusivity and unbalanced representation (an internal crisis), and to achieve *both* greater viability *and* social impact. This process of restructuring, which is intended to more comprehensively mirror the diversity and inclusivity of society, will unavoidably last years into the future, making this research necessarily a work-in-progress. To date, theatres and professionals have not neglected the potential of digital theatre in this process, but nor have they exploited hybrid drama or technology as much as they did in the first year of the pandemic. The 2022 revivals, the modern staging of classics, the debuts of new plays, and the performances that were transferred from off-Broadway to Broadway, or from non-West End to West End theatres, or from US to UK theatres and vice versa, did show greater inclusivity of authors from traditionally marginalised social groups (women, BIPOC and/or LGBTQ+). This tendency has also been reflected in the diversity displayed in the US and UK “best of” lists. A final observation that can be made in the conclusion to this paper is that the recent activity in the field is similar to that of the first two years of the pandemic, in that most playtexts and performances given in the first post-pandemic season steered clear of the historic event of COVID-19, both in its literal/symbolic sense, and its application. The one play that has woven the pandemic into its dramatic story and plot in a substantial way is Alecky Blythe’s *Our Generation*, whose playtext was assessed in this paper.

## References

- Akbar, Arifa. 2022a. “Daddy: A Melodrama review – Black art, white money and love in Speedos.” *The Guardian*, April 7, 2022. <https://www.theguardian.com/stage/2022/apr/07/daddy-a-melodrama-review-black-art-white-money-almeida-london>.
- . 2022b. “Prima Facie review – Jodie Comer on formidable form in roaring drama.” *The Guardian*, April 28, 2022. <https://www.theguardian.com/stage/2022/apr/27/prima-facie-review-jodie-comer-on-formidable-form-in-roaring-drama>.
- . 2022c. “Two Palestinians Go Dogging review – devastatingly human portrait of conflict sears itself on the mind.” *The Guardian*, May 23, 2022. <https://www.theguardian.com/stage/2022/may/23/two-palestinians-go-dogging-review-royal-court-sami-ibrahim>.
- . 2022d. “The P Word review – an irresistible romance and so much more.” *The Guardian*, September 16, 2022. <https://www.theguardian.com/stage/2022/sep/16/the-p-word-review-an-irresistible-romance-bush-theatre>.
- . 2022e. “The best theatre of 2022.” *The Guardian*, December 15, 2022. <https://www.theguardian.com/stage/2022/dec/15/the-best-theatre-of-2022>.
- Akhtar, Waleed. 2022. “Voicing concern: When are accents not appropriate in theatre?” *The Stage*, June 14, 2022. <https://www.thestage.co.uk/opinion/voicing-concern-when-are-accents-not-appropriate-in-theatre>.
- Almeida Theatre. 2022. “Daddy: A Melodrama.” *Almeida Theatre*, March 26–April 30, 2022. <https://almeida.co.uk/whats-on/daddy-a-melodrama/>.
- Anderson, Lisa M. 2008. *Black Feminism in Contemporary Drama*. Champaign: University of Illinois Press.

- Angelaki, Vicky. 2017. *Social and Political Theatre in the 21st Century Britain: Staging Crisis*. London: Bloomsbury Methuen Drama.
- Armistead, Claire. 2022. "The Father and the Assassin Review – Gripping tale of the man who killed Gandhi." *The Guardian*, May 20, 2022. <https://www.theguardian.com/stage/2022/may/20/the-father-and-the-assassin-review-national-theatre>.
- Atlantic Theatre. 2022. "The Bed Wetter." *Atlantic Theatre*. <https://atlantictheater.org/production/the-bedwetter/>.
- BBC News. 2023. "What is Windrush and who are the Windrush generation?" *BBC*, July 27, 2023. <https://www.bbc.com/news/uk-43782241>.
- Blake, Elissa. 2020. "Zoom fatigue and an awkward in-between: theatre can't wait to be back onstage." *The Guardian*, June 15, 2020. <https://www.theguardian.com/stage/2020/jun/16/zoom-fatigue-and-an-awkward-in-between-theatre-cant-wait-to-be-back-onstage>.
- Blythe, Alecky. 2022. *Our Generation*. London: Nick Hern Books.
- Broadway League. 2022. "2022 – 75th Annual Tony Awards." *Tony Awards*. Accessed December 15, 2023. <https://www.tonyawards.com/history/year-by-year/2022/>.
- Canning, Charlotte M. 2020. "Theatre and the last pandemic." *American Theatre (Publication of TCG)*, March 24, 2020. <https://www.americantheatre.org/2020/03/24/theatre-and-the-last-pandemic/>.
- Cavendish, Dominik. 2022. "The best theatre of 2022." *The Telegraph*, December 29, 2022. <https://www.telegraph.co.uk/theatre/what-to-see/best-theatre-plays-2022/>.
- Čirić-Fazlija, Ifeta. 2022. "Staging race and gender in the era of contemporary crises: Dramas of African American women playwrights." *American, British and Canadian Studies* 39 (1): 54–78. <https://doi.org/10.2478/abcsj-2022-0017>.
- Clapp, Susannah. 2022. "Theatre: Susannah Clapp's 10 best shows of 2022." *The Guardian Observer*, December 18, 2022. <https://www.theguardian.com/stage/2022/dec/18/susannah-clapp-10-best-theatre-shows-of-2022>.
- Colleary, Eric. 2021. "Theatre 2020 Project: Documenting a year of changes." *Harry Ransom Centre, U of Texas at Austin, YouTube Channel*. [https://www.youtube.com/watch?v=sKEg2Wewmmk&list=PL8cM00nJmOyf0WHgyz0RQb48KHOe2O\\_cl&index=1&t=381s](https://www.youtube.com/watch?v=sKEg2Wewmmk&list=PL8cM00nJmOyf0WHgyz0RQb48KHOe2O_cl&index=1&t=381s).
- Concord Theatricals. 2023. "The Life." *Concord Theatricals*. <https://www.concordtheatricals.com/p/44645/the-life>.
- . 2022. "Ohio state murders: Details". *Concord Theatricals*. <https://www.concordtheatricals.com/p/2569/ohio-state-murders>.
- Drama Online Library. 2023. "Alecky Blythe." *Drama Online Library*. [https://www.dramaonlinelibrary.com/person?docid=person\\_blytheAlecky](https://www.dramaonlinelibrary.com/person?docid=person_blytheAlecky).
- Dramatists Play Service. 2009. "Cock." *Dramatists Play Service*. <https://www.dramatists.com/cgi-bin/db/single.asp?key=4555>.
- Farber, Jim. 2023. "The wake-up call: Performing arts face the post-pandemic world." *San Francisco Classical Voice*, May 2, 2023. <https://www.sfcv.org/articles/feature/wake-call-performing-arts-face-post-pandemic-world>.
- Feldman, Adam. 2022. "The top 10 NYC theater productions of 2022." *Time Out*, December 22, 2022. <https://www.timeout.com/newyork/news/the-top-10-nyc-theater-productions-of-2022-122222>.
- Fernández-Caparrós, Ana, and Anna M. Brígido-Corachán. 2017. "Introduction: Re/presentations of crisis in twenty-first-century US literature and culture." *Studies in the Literary Imagination* 50 (1): v–xx. <https://doi.org/10.1353/sli.2017.0002>.
- Fisher, Philip. 2006. "Cruising." *British Theatre Guide*. <https://www.britishtheatreguide.info/reviews/cruising-rev>.
- Fonner, Daniel F., Zanie Giraud Voss, Glen B. Voss, Teresa Eyring, Rachael Hip-Flores, and Corinna Schulenberg. 2022. "Theatre facts 2021: Theatre Communication Group's report on the fiscal state of the U.S. professional not-for-profit theatre field." *Theater Communication Group*. <https://tcg.org/Web/Web/Our-Work/Research/Theatre-Facts.aspx>; <https://docs.google.com/document/d/1Co-d3NPBOiUfx7bvgTOAydFVvxtHGCNeB2bbAUReYjo/edit?usp=sharing>.

- Frank, Glenda. 2022. "Take Me Out: Helen Hayes Theatre." *Plays International & Europe*, April 17, 2022. <https://playsinternational.org.uk/take-me-out-helen-hayes-theatre/>.
- Gelt, Jessica. 2022. "Theatre is back but its recovery is proving to be perilous." *LA Times*, May 11, 2022. <https://www.latimes.com/entertainment-arts/story/2022-05-11/theater-is-back-attendance-is-down-costs-soaring-covid-19-pandemic>.
- Gordon, David. 2022. "The 10 best Broadway shows of 2022." *Theatermania*, December 27, 2022. [https://www.theatermania.com/news/the-10-best-broadway-shows-of-2022\\_94601/](https://www.theatermania.com/news/the-10-best-broadway-shows-of-2022_94601/).
- Green, Jesse. 2022a. "Review: 'Wedding Band,' a searing look at miscegenation nation." *The New York Times*, May 8, 2022. <https://www.nytimes.com/2022/05/08/theater/review-wedding-band.html>.
- 2022b. Review: Sarah Silverman's 'Bedwetter' musical has sprung a leak." *The New York Times*, June 7, 2022. <https://www.nytimes.com/2022/06/07/theater/review-sarah-silvermans-bedwetter.html>.
- Green, Jesse, Maya Phillips, Laura Collins-Hughes, Scott Heller, Alexis Soloski, and Elisabeth Vincentelli. 2022. "Best theater of 2022." *The New York Times*, December 5, 2022. <https://www.nytimes.com/2022/12/05/theater/best-theater-broadway.html>.
- Hall, Margaret. 2023. "Black boys who have considered suicide when the hue gets too heavy to play West End." *Playbill*, January 19, 2023. <https://playbill.com/article/black-boys-who-have-considered-suicide-when-the-hue-gets-too-heavy-to-play-West-End>.
- . 2022. "Cinco Paul and Bekah Brunstetter musical A.D. 16. streams March 16." *Playbill*, March 16, 2022. <https://playbill.com/article/cinco-paul-and-bekah-brunstetter-musical-a.d.-16-streams-march-16>.
- Hazemali, David, and Tomaž Onič. 2023. "'Canning the Keiser' in words and images: Case studies of patriotic American propaganda from WWI." *ELOPE: English Language Overseas Perspectives and Enquiries* 20 (2): 49–63. <https://doi.org/10.4312/elope.20.2.49-63>.
- Hemming, Sarah. 2022. "Best theatre shows of 2022." *The Financial Times*, December 28, 2022. <https://www.ft.com/content/14b8dc9d-bafa-41bf-b0a7-6f6ce6c1ff48>.
- Hempkin, Kirsten. 2016. "Exploring student attitudes to the refugee crises: Songs on migration." *ELOPE: English Language Overseas Perspectives and Enquiries* 13 (1): 111–21. <https://doi.org/10.4312/elope.13.1.111-121>.
- Huston, Caitlin. 2022. "Jordan E. Cooper on saving Broadway's 'Ain't No' Mo,' his 'Love and Loath' letter to American theater." *The Hollywood Reporter*, December 12, 2022. <https://www.hollywoodreporter.com/lifestyle/arts/aint-no-mo-jordan-e-cooper-saving-broadway-play-1235280568/>.
- Kennedy, Mark. 2022. "Broadway musical 'Kimberly Akimbo' creates upside-down world." *The Independent*, December 19, 2022. <https://www.independent.co.uk/news/ap-new-york-broadway-pulitzer-prize-mark-kennedy-b2248133.html>.
- Kousoulis, Antonis A., Konsantinos P. Economopoulos, Effie Poulakou-Rebelakou, George Androutsos, and Sotirios Tsiodras. 2012. "The plague of Thebes, a historical epidemic in Sophocles' Oedipus Rex." *Emerging Infectious Diseases* 18 (1): 153–57. <https://doi.org/10.3201/eid1801.AD1801>.
- Lambie, Sarah. 2022. "Meet the MDEE speaker: Alecky Blythe." *Drama and Theatre*, December 1, 2022. <https://www.dramaandtheatre.co.uk/features/article/meet-the-mdee-speaker-alecky-blythe>.
- Langston, Stephen. 2022. "The pandemic nearly killed theatre – The creative way it fought back could leave it stronger." *The Conversation*, February 21, 2022. <https://theconversation.com/the-pandemic-nearly-killed-theatre-the-creative-way-it-fought-back-could-leave-it-stronger-176185>.
- Lewis, Christian. 2019. "Review: Jeremy O. Harris's 'Daddy' is a masterpiece of Melodrama." *Medium*, March 6, 2019. <https://clewisreviews.medium.com/review-jeremy-o-harriss-daddy-is-a-paternal-masterpiece-b95fb9f8a058>.
- Llyr Evans, Gareth. 2022. "Celebrated Virgins review – a loving tribute to the Ladies of Llangollen." *The Guardian*, May 26, 2022. <https://www.theguardian.com/stage/2022/may/26/celebrated-virgins-review-ladies-of-llangollen>.
- Lukowski, Andrzej. 2022. "The best London theatre shows of 2022." *Time Out*, December 9, 2022. <https://www.timeout.com/london/theatre/the-best-london-theatre-shows-of-2022>.

- Lunden, Jeff. 2020a. "The show must go online: Theaters closed by COVID-19 get creative." *NPR*, April 11, 2020. <https://www.npr.org/2020/04/11/830390452/the-show-must-go-online-theaters-closed-by-covid-19-get-creative>.
- . 2020b. "Theaters return to an old art form – the radio drama – with a twist." *NPR*, June 9, 2020. <https://www.npr.org/2020/06/09/865061847/theaters-return-to-an-old-art-form-the-radio-drama-with-a-twist>.
- Marks, Peter. 2022a. "Saints be praised: Mary Magdalene and Jesus are meet-cute Nazareth teens in a hilarious 'A.D. 16'." *The Washington Post*, February 16, 2022. <https://www.washingtonpost.com/theater-dance/2022/02/16/ad-16-musical-olney-theatre-mary-magdalene/>.
- . 2022b. "The best of theater in 2022." *The Washington Post*, December 13, 2022. <https://www.washingtonpost.com/theater-dance/2022/12/13/best-theater-2022-broadway-london/>.
- Marlowe, Sam. 2022. "Top 50 shows of 2022: London." *The Stage*, December 16, 2022. <https://www.thestage.co.uk/features/top-50-shows-of-2022-london-age-of-vengeance-at-the-barbican-iphigenia-in-splott-at-lyric-hammersmith-theatre-the-crucible-at-the-national-theatre>.
- Mason, Amelia. 2023. "Boston theatres survived COVID. Can they survive what's next?" *WBUR*, May 1, 2023. <https://www.wbur.org/news/2023/05/01/boston-theater-pandemic-whats-next>.
- McHenry, Jackson. 2022. "The Best Theatre of 2022." *Vulture*, December 9, 2022. <https://www.vulture.com/article/best-theater-2022.html>.
- McKernan, Bethan, and Sufian Taha. 2023. "Shireen Abu Aqleh: 'Cold-blooded' killing and funeral chaos leave West Bank in turmoil." *The Guardian*, May 15, 2023. <https://www.theguardian.com/world/2022/may/15/cold-blooded-killing-and-funeral-chaos-leave-west-bank-in-turmoil>.
- McNulty, Charles. 2022a. "Review: Tony-winning 'The Inheritance' is better in Los Angeles than on Broadway." *LA Times*, October 18, 2022. <https://www.latimes.com/entertainment-arts/story/2022-10-18/review-matthew-lopezs-tony-winning-the-inheritance-gains-strength-in-los-angeles>.
- . 2022b. "Best theater of 2022: So many L.A. bright spots in a challenging year." *LA Times*, December 4, 2022. <https://www.latimes.com/entertainment-arts/story/2022-12-04/best-2022-theater-inheritance-slave-play-oklahoma>.
- Nottage, Lynn. 2022. "From Lynn Nottage: The true story behind *Intimate Apparel*." *PBS (web page)*, October 5, 2022. <https://www.pbs.org/wnet/gperf/blog/from-lynn-nottage-the-true-story-behind-intimate-apparel/>.
- Ogawa, Aja. n.d. "Projects." *Aja Ogawa Web Page*. Accessed May 15, 2023. <https://ayaogawa.com/projects#/the-nosebleed>.
- Oladipo, Gloria. 2022. "Ain't No Mo review – Absurdist and uneven satire about race in America." *The Guardian*, December 5, 2022. <https://www.theguardian.com/stage/2022/dec/05/aint-no-mo-review-broadway-jordan-e-cooper>.
- Paulson, Michael, and Javier C Hernández. 2022. "Live performance is back. But audiences have been slow to return." *The New York Times*, August 21, 2022. <https://www.nytimes.com/2022/08/21/arts/performing-arts-pandemic-attendance.html>.
- Philips, Maya. 2022. "10 stages and screens where I saw connection." *New York Times*, December 6, 2022. <https://www.nytimes.com/2022/12/06/theater/theater-performance-tv-2022.html>.
- Playbill. 2023. "The Life." *Playbill*. <https://playbill.com/production/the-life-ethel-barrymore-theatre-vault-0000004302>.
- Plemenitaš, Katja. 2016. "Songs as elements in the generic structure of film musicals." *ELOPE: English Language Overseas Perspectives and Enquiries* 13 (1): 31–38. <https://doi.org/10.4312/elope.13.1.31-38>.
- Pulitzer Prizes. 2002. "*Topdog/Underdog*, by Suzan-Lori Parks." *Pulitzer Prizes.org* <https://www.pulitzer.org/winners/suzan-lori-parks>.
- . 1998. "How I Learnt to Drive, by Paula Vogel." *Pulitzer Prizes.org*. <https://www.pulitzer.org/winners/paula-vogel>.
- . 2020. "*A Strange Loop*, by Michael R. Jackson." *Pulitzer Prizes.org*. <https://www.pulitzer.org/winners/michael-r-jackson>.

- Ryan, Anya. 2023. "For Black Boys ... review – A moving meditation on Black masculinity." *The Guardian*, April 2, 2023. <https://www.theguardian.com/stage/2023/apr/02/for-black-boys-review-apollo-theatre-london>.
- Shakespeare, William. 1982. "Romeo and Juliet." In *The Illustrated Stratford Shakespeare*, 701–28. London: Chancellor Press.
- Shaw, Helen. 2022. "A speed date with *Long Day's Journey into Night*." *Vulture*, January 25, 2022. <https://www.vulture.com/2022/01/theater-review-long-days-journey-into-night-elizabeth-marvel.html>.
- ShowTown Theatricals. 2022. "American Prophet: Frederick Douglass in his own words." *ShowTown Theatricals*. <https://www.americanprophetthemusical.com/>.
- Sierz, Aleks. 2019. "Helen Edmundson's 'Small Island' at the National Theatre: Tribute to the Windrush generation." *The Theatre Times*, May 9, 2019. <https://thetheatretimes.com/helen-edmundsons-small-island-at-the-national-theatre-tribute-to-the-windrush-generation/>.
- Smith, David. 2022. "Frederick Douglass gets the Hamilton treatment with new musical." *The Guardian*, July 30, 2022. <https://www.theguardian.com/stage/2022/jul/30/frederick-douglass-musical-american-prophet>.
- SMU DataArts. 2023. "Reduced expenses and growth in government support kept theatres afloat in 2021." *SMU Data Arts*, February 07, 2023. <https://culturaldata.org/learn/data-at-work/2023/tcg-theatre-facts-2021/>.
- Star Insights. 2022. "LA's performing arts and reopening survey: Overview of survey findings." *Star Insights*, February 10, 2022. [https://www.artsforla.org/wp-content/uploads/2023/07/LAs\\_Performing\\_Arts\\_\\_\\_Reopening\\_-\\_Survey\\_Learnings\\_Feb2022.pdf](https://www.artsforla.org/wp-content/uploads/2023/07/LAs_Performing_Arts___Reopening_-_Survey_Learnings_Feb2022.pdf).
- Stephens, Simon. n.d. "S2 Ep13: Alecky Blythe talks to Simon Stephens." *Royal Court Theatre Podcast*. Accessed December 15, 2023. <https://royalcourttheatre.com/podcast/s2-ep13-alecky-blythe-talks-simon-stephens/>.
- The Old Vic. 2022. "A Number." *The Old Vic*, January 24–March 19, 2022. <https://www.oldvictheatre.com/stage/event/a-number>.
- Tichenor, Austin. 2021. "Speaking what we feel: Shakespeare's plague plays." *Folger Shakespeare Library*. <https://www.folger.edu/blogs/shakespeare-and-beyond/speaking-what-we-feel-shakespeares-plague-plays/>.
- Towers, Andrea. 2022. "Take Me Out review: A star-studded cast takes center field on Broadway." *Entertainment Weekly*, October 26, 2022. <https://ew.com/theater/take-me-out-broadway-review/>.
- Travers, Peter. 2019. "'The Inheritance' is a ravishing theatrical work that urges generations to connect and love." *Rolling Stone*, November 18, 2019. <https://www.rollingstone.com/culture/culture-news/the-inheritance-play-broadway-review-913859/>.
- United Agents. n.d. "Alecky Blythe: Writer." *United Agents*. Accessed December 15, 2023. <https://www.unitedagents.co.uk/alecky-blythe>.
- Veltman, Chloe. 2023. "Theater never recovered from COVID – and now change is no longer a choice." *NPR*, February 6, 2023. <https://www.npr.org/2023/02/06/1153453450/oregon-shakespeare-control-group-productions-west-village-co-op-theater>.
- Watford, Daisy. 2022. "16 great shows to see in 2022 in London's most beautiful theatres." *Matadornetwork*, January 6, 2022. <https://matadornetwork.com/read/best-london-shows-2022/>.
- Weinert-Kendt, Rob. 2023. "If you rebuild it, will they return?" *American Theatre (Publication of TCG)*, March 20, 2023. <https://www.americantheatre.org/2023/03/20/if-you-rebuild-it-will-they-return/>.
- Wiegand, Chris. 2022. "Tony Awards 2022: A Strange Loop leads the nominees, while British talents are lauded." *The Guardian*, May 9, 2022. <https://www.theguardian.com/stage/2022/may/09/tony-awards-2022-nominees-strange-loop-the-lehman-trilogy-six-mj>.
- Willcocks, Geoff. 2008. "Europe in flux: Exploring revolution and migration in British plays of the 1990." In *A Concise Companion to Contemporary British and Irish Drama*, edited by Nadine Holdsworth and Mary Luckhurst, 7–25. Malden: Blackwell Publishing.
- Williams, Gareth. 2022. "5 minutes with Katie Elin-Salt, writer of 'Celebrated Virgins', Theatr Clwyd." *Get The Chance*, May 19, 2022. <https://getthechance.wales/2022/05/19/5-minutes-with-katie-elin-salt-writer-of-celebrated-virgins-theatr-clwyd/>.

- Williams, Holly, Jessie Thompson, and Isobel Lewis. 2022. "Jodie Comer, a gen-Z Elle Woods and a giant swimming pool: The best theatre of 2022, ranked." *The Independent*, December 19, 2022. <https://www.independent.co.uk/arts-entertainment/theatre-dance/features/best-theatre-2022-london-b2246017.html>.
- Wolley, Agnes. 2014. *Contemporary Asylum Narratives: Representing Refugees in the 21st Century*. London: Palgrave Macmillan.
- Woolman, Natalie. 2012. "Alecky Blythe: 'To get interviews there are certain risks'." *The Observer*, April 8, 2012. <https://www.theguardian.com/theobserver/2012/apr/08/alecky-blythe-verbatim-theatre-interview>.
- Young, Robin, and Allison Hagan. 2020. "'He didn't flee': Shakespeare and the plague." *WBUR*. <https://www.wbur.org/hereandnow/2020/04/06/shakespeare-plague-coronavirus>.

Vesna Tripković-Samardžić  
University "Mediterranean" Podgorica,  
Montenegro

2024, Vol. 21 (1), 265-282(316)  
journals.uni-lj.si/elope  
<https://doi.org/10.4312/elope.21.1.265-282>  
UDC: 821.111+  
821.111(73)-2.09:792(497.16)"2020/..."

## Contemporary Anglophone Plays on Montenegrin Stages in the Post-Pandemic Era

### ABSTRACT

This paper explores the reception and impact of nine contemporary Anglophone plays in Montenegro during the post-pandemic era. Analysing data from theatre archives, reviews, and interviews, the study examines the selection, production, and reception of these plays in the context of Montenegro's political, social, economic, and cultural landscape. The research highlights a sustained interest in both classic and modern-day Anglophone playwrights and delves into complex societal challenges, providing a backdrop for understanding the theatre's role in addressing these issues. The adapted plays, which explore global concerns and existential themes, contribute to a dialogue with Montenegrin society and challenge accepted norms. The absence of COVID-related plays is discussed, revealing programming policies and priorities. Ultimately, this paper explores how theatre continues to be a vital platform for critique and engagement with contemporary issues, reaffirming its role as a space for dialogue and collective therapy.

**Keywords:** Anglophone plays, post-pandemic, theatre, Montenegro

## Sodobna anglofonska dramska dela na črnogorskih odrih v obdobju po pandemiji

### IZVLEČEK

Avtorica obravnava recepcijo in vpliv devetih sodobnih anglofonskih dramskih del v Črni gori v obdobju po pandemiji. S pomočjo podatkov iz gledaliških arhivov, kritik in intervjujev preučuje izbor, produkcijo in recepcijo teh iger v črnogorski politični, družbeni, gospodarski in kulturni krajini. Izsledki kažejo, da se uprizarjata tako klasična kot sodobna anglofonska dramska dela, pri čemer so prikazane tudi zapletene družbene spremembe, ki pomagajo razumeti vlogo gledališča pri reševanju teh vprašanj. Dramske priredbe, ki se ukvarjajo s svetovnimi izzivi in vprašanji človeškega obstoja, vzpodbujajo dialog s črnogorsko družbo in prevprašujejo uveljavljene družbene norme. Avtorica se sprašuje, zakaj se na repertoarjih ne pojavljajo dramske igre, povezane z virusom COVID, kar kaže na programske usmeritve in interes gledališč. Obravnavano je tudi vprašanje, v kakšni meri je gledališče še vedno pomembna platforma za kritiko in ukvarjanje s sodobnimi vprašanji, pri čemer avtorica ugotavlja, da je gledališče še vedno prostor za dialog in kolektivno terapijo.

**Ključne besede:** anglofonske dramske igre, obdobje po pandemiji, gledališče, Črna gora

# 1 Introduction

Theatres worldwide observed World Theatre Day on March 27, 2020 without an audience. COVID-19 forced the curtains to fall, prompting theatres to transition to virtual environments in order to sustain the artistic spirit and alleviate the period of required isolation in the battle against the coronavirus.

The first anti-pandemic measures related to theatres in Montenegro were implemented on March 12, 2020. The theatres, along with cultural institutions engaged in theatrical production and numerous festivals functioning as institutions,<sup>1</sup> suspended their regular repertoires. “The context closed theatres”, stated one of the leading Montenegrin teatrologists, Janko Ljumović (2021, 4), such that theatre as “a living art form and the performance as an act that unfolds between performers and the audience became an impossible mission”. The crisis brought to light the fragility of live theatrical art, with risks to artistic production processes, including the impossibility of participating in festivals and uncertainties surrounding planned premieres before and during the pandemic (Ljumović 2023, 185). As the situation unfolded the theatres in Montenegro waited a long time for protocols that would allow them to reopen, mainly due to the fact that “the cultural system was not a priority”, but also because “hardly any voice seeking possibilities was heard” (Ljumović 2021, 4). When the protocols finally appeared, they established maximum numbers of audience members and daily performances, to which theatres had difficulty adjusting.

Nevertheless, theatres and cultural institutions in Montenegro tried to respond creatively to the crisis. Physical distance, the “new normal” (Ljumović 2021, 4), opened the need for social connections, a secure space for that being the internet. Some experts in the dramatic arts argued that theatre had proven to be vital and adaptable to the new conditions, while others believed it should strictly adhere to its established norms (Marojević 2020, 21). It was at least shown that, despite the crisis, and perhaps because of it, theatre remained indispensable. We shall consider the activities of four institutions significant in this context, especially with regard to their Anglophone productions.

During the pandemic, the Montenegrin National Theatre adapted to challenges through the multimedia project *Živimo kulturu* (living culture), supported by the Ministry of Culture and implemented through various media channels (Ljumović 2023, 185). Collaborating with Radio and Television of Montenegro, the theatre produced *Poetski teatar* (poetic theatre), indicating a shift towards establishing a “multimedia theatre” (Marojević 2020, 33). Importantly, amid the prevailing restrictive conditions of the 2020/2021 season, two premieres (both of them Anglophone plays!) were successfully staged: Viktor Lodato’s *The Woman Who Amuses Herself* (*Žena koja se igrat Đokonda*), and Lucy Prebble’s *The Effect*.

---

<sup>1</sup> In Montenegro, a small country with approximately 630,000 people, the focal points of theatre production comprise two national theatres – the Montenegrin National Theatre and the Royal Theatre Zetski dom – along with two city theatres, the City Theatre Podgorica and Nikšić Theatre. Furthermore, there are irregular and sporadic theatrical productions organized by municipal cultural centres and summer theatre festivals. The independent theatre scene is characterized by underdevelopment and operates in a significantly unequal position compared to budgetary institutions.

In this period there were still no digital transformation strategies, and the Ministry of Culture's only contribution was the archival recordings of performances broadcast on Radio and Television of Montenegro (Ljumović 2021, 4). One exception to this trend was the unique initiative of the Royal Theatre Zetski dom entitled *Zetski dom sa Vama* (Zetski dom with you), which unfolded on social media during the early months of the pandemic. Diverse themes brought together friends, collaborators, and the audience of this theatre through short video narratives, stories, and experiences, emphasizing the foundational role of dialogue and communication in theatre (Ljumović 2021, 5). However, the premiere and reruns of the Anglophone play *Who's Afraid of Virginia Woolf?* (*Ko se boji Virdžinije Vulf?*), directed by Paolo Magelli, initially scheduled for October 7, 2020 at the same theatre, were postponed due to the coronavirus.

Amidst the challenges posed by the pandemic, the City Theatre Podgorica adapted to the circumstances, bringing joy to children through various online initiatives (through social media, a website, and YouTube channel) including workshops, plays, songs, and stories (Marojević 2020, 21–22). This theatre initiated virtual programs on the World Day of Theatre for Children and Young People, successfully launched the Regional Internet Theatre Festival for Children, staged one premiere, and initiated rehearsals and preparations for two other plays. Despite the lack of performances, Nikšić Theatre maintained an active presence through intense publishing activities, releasing one double and one standard issue of its journal *Pozorište* (*Theatre*) online, which chronicled events in other theatres through various media channels, such as TV and radio.

Transitioning from the historical backdrop of Montenegrin theatres' responses to the pandemic and its aftermath, the subsequent subsections examine the reception of nine contemporary Anglophone plays staged in Montenegro in the years since.

Within the framework of this paper, the term “contemporary” refers to plays that emerged from the 1950s to the year 2016. The productions in Montenegro were based on new Montenegrin translations or Serbian translations, some of which had been used in well-known theatre productions. The translations remained largely faithful to the original English texts. The plays were staged in four theatres (the Royal Theatre Zetski dom, Cetinje, the Montenegrin National Theatre, Podgorica, Nikšić Theatre, and the City Theatre Podgorica) in the post-pandemic period, from 2021 to 2023. We have chosen to research the stage productions of these theatres because they form the focal points of theatre production in Montenegro. Besides, these are the only professional theatres in Montenegro that function as public institutions: the Montenegrin National Theatre and the Royal Theatre Zetski dom have the status of national theatres, and the other two are city theatres. Further, the Royal Theatre Zetski dom and Nikšić Theatre share the longest theatre tradition in Montenegro, dating back to the year 1884.

The first section highlights Montenegro's enduring interest in adapting Anglophone plays, and explores the significance of contemporary Anglophone plays staged in Montenegrin theatres over the past three years, examining their relevance to global issues, the political contextualization of the choices of the plays staged, the role of the plays in challenging local societal norms, and their contribution to redefining the purpose of theatre amidst modern challenges.

The second section offers an in-depth study of the reception of each contemporary Anglophone play through the lens of theatre archives, critical reviews, and interviews with theatre practitioners, examining key themes, translation challenges, production details, and reception by both critics and audiences. The chapter sheds light on the plays' resonance within the Montenegrin context, providing the essential social, political, and theatrical context for a deeper understanding of each play. The analysis is structured around specific theatres, not only for clarity but also to reveal the nuanced selection process aligned with each theatre's programming policy.

## **2 Contemporary Anglophone Plays on Montenegrin Stages: An Overview and Socio-Political Contextualization**

The theatre and performance archives in Montenegro reveal a longstanding interest in contemporary Anglophone playwrights, dating back to the 1950s (Tripković-Samardžić 2019, 220). Before 2000, Montenegrin theatre practitioners exhibited a keen interest in adapting not just earlier classic playwrights such as William Shakespeare and Ben Jonson, but also prominent literary figures of the twentieth century (Tripković-Samardžić 2019, 220). Research on the reception of contemporary Anglophone drama in Montenegrin theatres in the first two decades of the twenty-first century affirms the enduring appeal of "classic" playwrights among Montenegrin audiences, while concurrently witnessing an increasing fascination with adapting contemporary playwrights such as Edward Bond, Steven Berkoff, and Martin McDonagh (Tripković-Samardžić 2019, 219). An in-depth analysis of the reception of 24 contemporary Anglophone plays staged in Montenegrin theatres in this period revealed that the decisive factors contributing to the selection and positive reception of particular authors and their plays were high artistic value, popularity and current relevance, directorial skill, actors' sensibilities, and the appeal of specific theatre genres (Tripković-Samardžić 2019, 228). The study also showed that the Anglophone plays staged reflected a trend towards politicization in response to contemporary challenges.

This research on the reception of contemporary Anglophone drama in Montenegrin theatres in the post-pandemic period builds upon the aforementioned study. Drawing upon data sourced from theatre archives, reviews, and interviews with Montenegrin theatre practitioners and scholars reveals that over a post-pandemic period spanning three years (commencing from the 2020/2021 theatrical season), four professional theatres in Montenegro have staged nine contemporary Anglophone plays. This rate of three per year is twice that seen in the first two decades of the twenty-first century. The findings affirm the enduring popularity of "classic" Anglophone playwrights such as Edward Albee, David Mamet, and John Osborne among the Montenegrin audience. Concurrently, there is a consistent trend of increasing interest in adapting works by modern-day Anglophone playwrights during the post-pandemic era, including Victor Lodato, Lucy Prebble, Simon Stone, Martin McDonagh, Duncan Macmillan, and Anthony Neilson.

In his article "We Need to Go Back to the Beginning", the journalist and theatre critic Bojan Munjin (2022, 5) argues that, following the near-complete halt of theatre activities during the pandemic, which forced theatres to grapple with the question of their physical survival, they

must now confront a new war, a new threat to humanity. The author depicts this new war as “a poisonous and deadly mixture composed of digital civilization, an era of cheap spectacle, a world driven by devouring profit, a devastated environment in which we suffocate, and finally, a pessimistic sense that we have lost the meaning of our own lives” (Munjin 2022, 5). Examining Montenegro’s distinctive political, social, economic, and cultural landscape over the past three years, particularly in the aftermath of the global pandemic, only enhances our understanding of the prevailing threat.

Particular to this landscape is that, in addition to the pandemic itself, during that period a significant political shift occurred in Montenegro: in August 2020, the ruling party of 30 years lost the elections. This change happened because of the endemic corruption permeating all spheres of society, including the cultural system, described by the theatre critic Maja Mrđenović (2020) as “generally contaminated, weakened, distorted by nepotism, corruption, clientelism, chaotic, and arbitrary”. The transformation triggered a collective sense of upheaval, prompting anxieties about forfeited privileges and national identity. The new political elite, navigating both political change and a global pandemic, failed to meet public expectations, reinforcing existing polarization. Initial optimism waned as the status quo persisted, leading to widespread disillusionment and historically low voter turnout. The European Commission Montenegro 2023 Report (2023, 5–6) underscored enduring challenges, such as corruption and discrimination, highlighting grave concerns over femicide, misogyny, gender-based violence, and violence against children.

In this light, we believe it is not just the inclusion of nine contemporary Anglophone plays in the repertoires of Montenegrin theatres over the last three years that is significant, but the particular choices: *The Woman Who Amuses Herself*, *The Effect*, *Yerma*, *Sexual Perversity in Chicago*, *Who’s Afraid of Virginia Woolf?*, *The Pillowman*, *Look Back in Anger*, *The Lungs*, and *The Lying Kind*. These are plays that not only reflect pressing global concerns but also resonate within the specific Montenegrin political, social, economic, and cultural context. Thus they contributed to the fostering of a dialogue with the Montenegrin community and the challenging of certain accepted notions underpinning a Montenegrin society in transition.

The plays, premiered in four Montenegrin theatres, address various pressing global issues, including environmental concerns (overpopulation, pollution, and animal cruelty), mental health (depression, the impact of antidepressants on brain chemistry, and dementia), fundamental human rights (women’s rights), and such societal issues as corruption, hypocrisy, class struggle, infertility, child abuse, paedophilia, and the commodification of sex. However, underlying these plays are existential themes such as confusion about the purpose of the survival of the human species, and the recognition of art, love, and fantasy as essential means of transcending the existing chaos.

The absence of contemporary Anglophone plays addressing the COVID pandemic, some of which introduce new or hybrid genres, is evident.<sup>2</sup> This can be attributed to the perception of the creators of programming policies in Montenegrin theatres, who may believe that Montenegrin audiences are not yet prepared to process the trauma of the pandemic. More

<sup>2</sup> The most recent instance of such plays is Suzan Lori Park’s Pulitzer Prize-winning *Plays for the Plague Year*.

significantly, it may also be that other global issues, or those that are more pertinent for the Montenegrin context, further intensified by the pandemic, have been considered more pressing. As such, these issues have demanded more attention from the theatre, a singular public space that, despite exclusion from traditional media, serves as a platform for freedom of speech. But then it may also be that after the pandemic theatre needed to go back to its roots, discern between virtue and madness, and communicate this message to the world (Munjin 2022, 5). In a chaotic world “that has decided to destroy itself”, Munjin states, the theatre faces novel challenges – “not only in how we discuss [the chaotic world] but also in determining what we should talk about and for what purpose” (5). What he suggests is that theatre should scrutinize fundamental values such as love, kindness, selflessness, and sacrifice, and seek the meaning of human existence, life, death, and suffering, regardless of how old-fashioned and conservative it may appear.

### 3 Reception of Nine Contemporary Anglophone Plays on Montenegrin Stages in the Post-Pandemic Era (2021–2023)

The examined plays belong to various genres (romance, dark thriller, melodrama, tragedy, slapstick, satire, black comedy) and time periods: from John Osborne’s “classic” *Look Back in Anger* (1956), through those published in the second half of the twentieth century – Edward Albee’s *Who’s Afraid of Virginia Woolf* (1962) and David Mamet’s *Sexual Perversity in Chicago* (1974) – to more recent ones, published in the twenty-first century: Anthony Neilson’s *The Lying Kind* (2002), Martin McDonagh’s *The Pillowman* (2003), Victor Lodato’s *The Woman Who Amuses Herself* (2011), *The Lungs* (2011), Lucy Prebble’s *The Effect* (2012) and Simon Stone’s *Yerma* (2016). In the following subsections, organized by hosting theatre, a more nuanced examination of the reception of these plays will be presented. This includes analysis of the themes addressed in the plays, production/adaptation challenges, theatre practitioners’ stances on the plays and the reception by critics and audiences.

#### 3.1 The Montenegrin National Theatre

Over the past three decades, the Montenegrin National Theatre has demonstrated its ability to offer new thematic interpretations of societal reality, as seen in numerous titles addressing transitional dilemmas and contradictions within the contemporary social context (Ljumović 2023, 186–87). During the period from 2021 to 2023, the Montenegrin National Theatre staged three contemporary Anglophone plays: two in the 2020/2021 season—*The Woman Who Amuses Herself*, which premiered on February 26, 2021, and *The Effect*, which premiered on June 9, 2021. Additionally, one play was staged in the 2021/2022 season: *Yerma* (*Jerma*) by Simon Stone, which premiered on December 29, 2022.

*The Woman Who Amuses Herself* (*Dokonda*) was the first performance to be staged in this theatre, almost a year after the lockdown. The play is based on a true story of Italian immigrant Vincenzo Peruggia, who, working as a glazier at The Louvre in 1911, stole the Mona Lisa and brought it to his bedroom, where he spent over two years captivated by his muse and her mysterious smile. He later attempted to sell it, and, despite the sensational crime, was also seen as a defender of Italy’s artistic heritage. Presented in fragments through monologue

scenes, the play raises significant questions that resonate with Montenegrin audiences: the helplessness and imperfection of man in the face of the perfection of art, the relationship between the common individual and art, the contrast between legal principles and personal conviction, the association of art with religious experience, inquiries into artistic heritage and the concept of patriotism, the intricate connection between man and art, the varied influences of art on individuals (including diverse interpretations and identifications with art objects), the role of art as consolation – a constant need for spirituality, beauty, and art during challenging times – and the role of a smile as a defence against death.

The director Dušanka Belada, who was also the dramaturge, notes that in an era where the aura of artistic works is diminishing due to hyperproduction and mass media, the play explores the reasons for fascination with a single artistic work from various perspectives, on multiple levels, and addresses the question of the meaning and function of art today (Marojević 2021c, 6). Although the play was the director's first professional production, it received unanimous acclaim from audiences, both at the premiere and during the subsequent performances. This is particularly noteworthy considering the director is a young Montenegrin, making her debut in the theatre world. In 2021, the play was featured at the Yugoslav Theatre Festival in Užice. On the official website of the Montenegrin National Theatre, in the news section related to this festival (Crnogorsko narodno pozorište, n.d.), the festival selector, Bojan Munjin, explains that the main reason this play resonates with all of us today is the fact that art is undervalued in the contemporary era of “the cheapest spectacle”, which adversely affects our attitudes to art.

The play also received an excellent reception from the theatre critics in Montenegro. Despite some dramatic shortcomings, such as the lack of focus in the first part of the play and caricatured secondary roles, the critics praised the production's use of lighting and music that contribute to the anti-realist discourse, as well as authentic theatrical character Vincenzo Peruggia, with whose futile struggle to give meaning to his life the audience empathizes, seeing the contrast between the perfection of a work of genius and human imperfection (Čukić Šoškić 2021a, 9–10). This effect was certainly achieved owing to the outstanding performance of the lead actor, Zoran Vujović, who found the play liberating as it underscores humanity's profound need for art and its noblest aspect – the smile – something we have all become lazy about. The critic Koprivica highlighted the director's fresh interpretation of the familiar text, using “new stage solutions with multimedia effects”, and praised her skill in balancing the text and stage action, and employing the “golden ratio of all stage parameters” (2021, 8).

Just three months after *Dokonda*, and in order to engage younger audiences by selecting a media language and themes that resonate with them, the Montenegrin National Theatre staged *The Effect* (2012). The play, written by an emerging British playwright and producer Lucy Prebble, falls between happy romance and dark thriller. It revolves around two characters, Tristan and Connie, who volunteer for a clinical trial of an antidepressant and unexpectedly fall in love. *The Effect* is both an intellectual and emotional exploration that delves into brain chemistry and its influences on our feelings and moods, examining the experience of being in love, and the causes of severe depression.

*The Effect* was a regional production, which holds particular significance for the National Theatre in Podgorica. The dramaturge Stefan Bošković emphasized that this play features a

complex and layered narrative that shifts our perspectives, intricately exploring the human mind, delving into brain investigation, pointing out medical limitations, probing into causes of depression, and measuring the power of physical attraction (Marojević 2021d, 6). As the protagonists continually rationalize their feelings, attempting to discern the authenticity of their emotions and the influence of brain chemistry, no solution is proposed, and instead the play leaves the audience with the realization that “the human brain remains an enigma for medicine, and our individual feelings elude rational explanations” (Čukić Šoškić 2021). Bošković embraced the challenge of adapting the play to Montenegro, recognizing substantial cultural differences from its original setting, and highlighted the demanding three-month process of tailoring the text to the Montenegrin context (Marojević 2021d, 6).

In an interview on January 20, 2024, the director Danilo Marunović stated that for this accessible and topical play, aimed at young audiences, one of the crucial aspects of adaptation was the language. As the director wanted audiences to connect with the play through language, he alone with the translator and dramaturge carefully handled the text (the translator S. Perović even attending the actual staging of the play), endeavouring to incorporate the Montenegrin dialect that was otherwise associated with stereotypical, archetypal characters. For the director, the pivotal question of the play – whether the love which arises due to a chemical reaction that has been induced in the brain by a drug is equivalent to the love which occurs naturally – is a theme that resonates with audiences in Montenegro as much as those in London. This is because, as the director concludes, we live in a pharmaceutical era, obsessed with illnesses, and have made it our identity.

Despite significant efforts in adapting the Montenegrin language, the director felt that the title should remain in its original form to preserve its ambiguity. The design, music, and choreography were designed for young audiences. As such, according to the director, audiences responded positively, even to the sexually explicit scenes, which, the director claimed, were presented in an innovative way. The copulation scene, for example, is first presented to the audience through an audio recording, intensifying the mutual seduction among the characters on stage, followed by a video presentation of chimpanzees mating – creating a visual contrast on the symbolic, physical, and choreographic levels. Nevertheless, due to external circumstances such as the COVID-19 pandemic and the main actress's inability to perform, the production ended early.

However, it appears that one of the reasons why the play received negative reviews, aside from a shortening of the text that lost some of the original text's semantic depth, was the use of a multitude of audio-visual effects that lacked expressive impact. The “unnecessary complexity of media-driven symbolisms” was one of the reasons the theatre critic Čukić Šoškić (2021b) labelled the play “theatrical/pretentious”. While Čukić Šoškić praised the set design for breaking the space into multiple levels and giving the play a modern touch, the critic specifically criticized the director's use of music – deeming it a mere auditory decoration with no connection to plot development or the characters' emotions – and excessive video materials to evoke animal instincts, which detracted from the impact of the actors on stage and diminished the overall effect.

The question of the purpose of art in the present-day era, as raised by the play *The Woman Who Amuses Herself*, is also addressed by theatre director, writer, and actor Simon Stone

(2018): “Why should people go to the theatre at all anymore? There are so many less time-, money- or effort-consuming ways to access narratives and entertainment. What is it that congregating in this ancient space brings to contemporary audiences?” Stone attempted to answer this question through his play *Yerma* (2016), based on Lorca’s tragic poem about the myth of longing for a child and the inability to conceive. He completely recontextualized the work, introducing a different kind of conflict, one between the political – a woman’s desire for autonomy in the face of societal expectations – and biological imperatives. The modern-day *Yerma*’s obsession with motherhood degrades her morally and psychologically, turning her life (marriage, family ties, and social status) into ruins. Attempting to appeal to contemporary audiences, Stone explained the effort he put into the staging of the play. With the deconstruction aiming to highlight the “documentary atmosphere in a theatrical setting” (2018) by challenging staged artificiality and fostering a transformative experience where audiences lose themselves in relatable individuals embodying classical theatre archetypes through inventive manipulations of space and time.

The play, produced at the Montenegrin National Theatre, marked the Belgrade director Ana Tomović’s first work in the country. Recognizing theatre as an art of the present moment, she emphasized the importance of *Yerma*, considering it a contemporary play with universal resonance, especially given its significance as a female-driven narrative. Tomović’s choice of work was also influenced by the significance of the play’s theme (the issue of ostracized women who cannot conceive) in the local Montenegrin (and broader Balkan) context. Furthermore, it reflects contemporary struggles and dilemmas faced by women (Četković 2022). The complexity of conveying these dilemmas on stage was confirmed by the lead actress, Kristina Obradović. For the actress, who grappled with the transformation of the protagonist’s evolving desire to conceive – from a casual idea through a genuine need to an obsessive struggle that fundamentally alters her life – the most challenging aspect of the character was attempting to bridge “the seemingly insurmountable gap between the idea of having no similarities with that woman and the eerie realization of how much of her is within [the actress]” (Višnjić 2023, 12).

The critics’ responses to the play were mixed. Čukić Šoškić (2023) praised the director for her subtle influence, unobtrusive set design, emphasis on the implosion of emotions in each scene through tense character relationships, and the shift of focus to the liberation of suppressed emotions at the play’s conclusion. Other critics, such as Nataša Nelević (2023), dismissed the director’s constraining ideological framework in exploring the concept of motherhood, echoing the criticism of Stone’s adaptation. Nelević argued that the contemporary *Yerma*’s desperate quest for motherhood and willingness to sacrifice everything to become a mother can only be viewed as abnormal and irrational, not tragic, reflecting twenty-first-century pragmatism and conformism that discard the imperative of female biology. Consequently, the adaptation was labelled a “psychological case study”, where the theme of motherhood is “interpreted as a pathological psychological phenomenon and subjected to criticism that is exclusively rationalistic and conformist”.

### 3.2 The Royal Theatre Zetski dom

The Royal Theatre Zetski dom is the oldest state theatre and cultural institution in Montenegro, and has established its dominance in the country’s cultural landscape through the successful

integration of transitional themes into a consistent programming strategy, skilfully avoiding traditional constraints (Ljumović 2019, 147). In line with this programming policy, the theatre produced two contemporary Anglophone plays in the 2021/2022 season: *Sexual Perversity in Chicago* (*Čikaške perverzije*) by David Mamet premiered on September 20, 2021; and *Who's Afraid of Virginia Woolf?* (*Ko se boji Virdžinije Vulf?*) by Edward Albee, whose premiere on October 7, 2020 was cancelled due to the pandemic, but was eventually staged on November 2, 2021. Additionally, one play was produced there in the season 2023/2024: *The Pillowman* (*Jastučko*) by Martin McDonagh, which premiered on September 17, 2023.

*Sexual Perversity in Chicago* (1974) by David Mamet is a dialogue-driven play featuring Bernie and Danny, urban males navigating a world of the 1970s where sexuality is commodified and fetishized. Through short, rapid scenes, they discuss their sexual fantasies and try to act them out with two women, Deb and Joan, who are also confused by contemporary sexual mores. The play highlights the alienation and the myth of consumption that surrounds sexuality. As stated by C.W.E. Bigsby (2001, 211), the play employs the singles bar as a metaphor for a society in which alienated individuals market themselves, and the play's humour arises from the characters' inability to comprehend themselves or others, creating an ironic gap between confident sexual language and awkward incompetence in real relationships.

The significance of this performance for the Royal Theatre and Montenegrin theatre more broadly is twofold. Firstly, it represents the inaugural staging of the play in Montenegro, filling a notable absence of the play on regional stages for an extended period. Simultaneously, it stands out as the debut performance featuring the complete ensemble of the newly formed acting troupe at the theatre. This Montenegrin production was based on a Serbian translation by Irena Kregar and Lary Zappia from 1991, which was used for the Yugoslav Drama Theatre's production in Belgrade in 1991, and later became a huge hit.

In the official programme for the performance, the director Rakočević (2021) explained that he drew inspiration from Mamet's response to the ideology of political-social idealism, depicting the story of an indoctrinated individual who is easily manipulated, which resonates with our reality. The director described Mamet's characters as lost, rejected, and frigid, stating that they "paranoically seek an escape, outside of themselves, beyond their own wreckage" (8) in a pornographized society that is devoid of content and does not instil hope, confidence, or motivation. According to Rakočević, Mamet's play proved to be much more inspiring than initially thought, leading the artists to present a somewhat different reading than expected with this play or the entire body of the author's work (Marojević 2021a, 13–14). The actor Vule Marković stated that the director and actors' new reading of the play resonates with the current moment, as it emphasizes the psychological perversions stemming from intimate frustrations and traumas, alienation, and the lack of communication in the present time (Marojević 2021a, 14). This is in line with how the actress Marija Labudović saw the world of the play – as depicting a world of technology where every perversion is accessible, overshadowing the importance of embrace, touch, and contact, especially with oneself, and the love we are all in desperate need of (Marojević 2021a, 13–14). This interpretation of the play is confirmed by the dramaturge Mišković, who saw the play as exploring "the castration of emotions resulting from overall transparency and [attempting] to break all taboos",

examining how we are losing ourselves in freedom, which paradoxically constrains us. The only solution for the four characters in emotional turmoil (of which they are unaware) turns out to be a simple thing: an embrace (Marojević 2021a, 13). Although the play resonated with Montenegrin audiences and offered an innovative interpretation, it was limited to only six performances with a lower-than-anticipated turnout (a total of just 272 people). The primary reason for this was the COVID-19 measures concerning audience capacity and seating arrangements.

Only a month after the premiere of *Sexual Perversity in Chicago*, the Royal Theatre staged another iconic twentieth-century Anglophone work, *Who's Afraid of Virginia Woolf?* (1962), a play that is widely celebrated for its detailed character study, intellectual cynicism, sado-masochism, and exploration of the moral and existential abyss. This conversation piece – static, like most of Albee's plays – explores a couple trapped in a toxic marriage made of lies and bitterness, and their decision to kill their fantasy child while forcing another couple to confront their own marital struggles. Here, Albee critiques American values, presenting an apocalyptic vision that depicts a struggle between the past and the future for the present (Bigsby 2001, 130). George, a history professor married to the daughter of the college president, symbolizes the past and rejects the new pragmatics embodied by Nick, a biologist representing the future. The play also examines the question of how we construct reality (Bigsby 2001, 130).

Even though the play reflects the societal implications of compensatory fictions and Albee's examination of the fate of American values, suggesting a deliberate erosion of communality and trust with broader cultural implications, the play resonates deeply with Montenegrin society. Due to the pandemic, the premiere of the play took place about a year later than initially planned and rehearsed for. The director Paolo Magelli, who had staged *Who's Afraid of Virginia Woolf?* in Dresden fifteen years before the premiere in Cetinje, did not find a direct connection to the previous production. Quoted in an article for the *Protagonist* magazine, entitled "Revolucija je začeta u teatru" (Revolution is Conceived in the Theatre; 2022, 6), while discussing the thematic and conceptual aspects of reading and analysing this play in Montenegro, Magelli stated that *Who's Afraid of Virginia Woolf?* is "not a literary but a political source", carrying revolutionary aspects and a quest for truth in both family and society: it begins with an analysis of how lies support society, criticizes the academic world, and challenges the broader societal norm that presents lies as the truth.

The translator and the dramaturge Željka Udovičić Pleština stated that the enduring relevance of *Who's Afraid of Virginia Woolf?* can also be attributed to the genre features of the play—its blend of dark comedy, melodrama, tragedy, slapstick, and satire. She also emphasized the contemporaneity of this work, which "lies in its fearless and unequivocal announcement of a penetrating, aggressive, and all-prepared 'new' generation that promotes the idea that success is measured solely by the criteria brought by liberal capitalism", and stressed the potential of the play to offer us the possibility of individual escape from societal conventions (Kontić 2021).

Yet, for the theatre critic Nataša Nelević (2021), the widespread popularity of the play raised the risk of diminishing its depth and reducing it to stereotypical interpretations. Nelević

argued that Magelli's reading of the play adheres closely to conventional interpretations, offering a familiar theatrical narrative that lacks substantial innovation despite incorporating visual spectacle, such as a giant chandelier on which the actors climb to share important words which are inaudible to the audience. While the critic praises the expressiveness of the actresses as well as the set design, the overall assessment highlights a deficiency in adopting novel approaches, including fresh theatrical language or new media, to address contemporary issues related to marital crises.

Regrettably, the production only ran for 12 performances. In a conversation in December 2023, the producer Aleksandra Maksimović revealed a notable challenge the production faced due to demands from the author's agency, which required approval for the director, actors, collaborators, set design, costumes, and textual alterations. Despite submitting all the required information to the agency, a disagreement between the theatre and agency arose over the inclusion of a chandelier in the set design (the agency requesting the chandelier's removal, despite having received sketches of it), ultimately resulting in the cancellation of the performance.

The latest play staged at this theatre is the contemporary Anglophone *The Pillowman* (2003), a black comedy by British-Irish playwright, film director, and screenwriter Martin McDonagh. The narrative revolves around Katurian, a fictional writer in an unidentified totalitarian state, interrogated by the detectives Ariel and Tupolski about the disturbing content of his stories, which resemble the bizarre murders of children that have been happening in his town. The play raises profound questions about art – the transcendental nature of art, the power of imagination, the interconnectedness between art and life, the artist's role in society, the cost of freedom of expression, the need for art (storytelling), and suffering as a source for creation. Additionally, it explores themes such as the contrast between innocence (childhood) and corruption (reality), along with controversial subjects like child abuse, emphasizing the imperative of openly addressing such issues. The play received the Laurence Olivier Award for Best New Play.

Despite being performed widely around the world, the play has only been staged three times in the region: twice in Serbia (Belgrade, Sombor) and once in Montenegro (Cetinje). The main reason the young Montenegrin director Andrija Rašović chose this text for his master's exam production was the fact that it celebrates the act of storytelling and “challenges its audience to confront the disturbing nature of stories, the power of imagination, and the delicate relationship of an unspoiled and innocent childhood faced with the harsh reality of the world we live in” (S. Vi 2023). In art, and especially in theatre, the director explained, “we should not shy away from the unspeakable. The central question we pose is: ‘What if, through stories, we give a voice to those who might otherwise remain silent?’” (S. Vi 2023).

The Montenegrin production was based on the Serbian translation made by Bojana Kovačević Petrović, which was then adapted to the Montenegrin language. In a brief conversation on January 11, 2024, the dramaturge Rajko Radulović mentioned that staging the story was challenging due to its complexity and unusual structure, making the text difficult to bring to life. However, he explained that the audience responded favorably to the play, as it helps them “to lose themselves in [it and] reevaluate views on life, art, [...] and the world around us, the

dark truths lying hidden beneath the surface of existence” (S. Vi 2023). The black humour in the play, as described by Radulović, is therapeutic – “a protective mechanism against the absurdity of existence” (S. Vi 2023). Given the recent production, at the time of writing this article the play had only been performed three times in Montenegro.

### 3.3 Nikšić Theatre

Meanwhile, Nikšić Theatre, which stands as the second oldest theatre in Montenegro, has premiered two contemporary Anglophone plays in the last three years: *Look Back in Anger* (*Osvrni se u gnjevu*) by John Osborne, which premiered on December 20, 2021, and *The Lungs* (*Pluća*) by Duncan Macmillan, which premiered on November 30, 2023. It was in line with its programming policy of adapting classics that the theatre opted for *Look Back in Anger* (1956), the seminal post-war British play which sparked the first wave of Kitchen-Sink Drama, giving voice to the frustrated and socially marginalized lower-middle class, whose literary icons were labelled the Angry Young Men. It is a 1950s domestic drama about the conflict between Jimmy Porter, an intelligent and educated young man from a working-class background, and his emotionally reserved upper-middle-class wife. On a deeper level, it is a play about the life and protest of the working class, about angry men and women who suffer, “not the old revolt of the proletariat against a tyrannical aristocracy, but rather the complaint of a frustrated lower-middle class against the failure of its overlords to define any code at all, around which the community could conduct a debate about who should inherit England” (Kiberd 2006, 29).

Despite the fact that the play reflects the societal upheavals in post-war England, its themes resonate profoundly with contemporary Montenegro, marked as it is by polarizations and divisions along class, national, and religious lines. The play explores various critiques, including the gap between social classes, the intertwining of the church with ruling elites, the hypocrisy among believers, and the apathy, ignorance, and false patriotism of the ruling classes, along with their neglect of the common man. Additionally, it emphasizes the pressing need for ideals, solidarity, audacity, sensitivity, and enthusiasm. All these reflect the challenges and complexities of Montenegro’s current milieu. For this reason, the Montenegrin audience could easily identify with the characters in this play and the conflicts it portrays.

Janko Jelić, the art director of Nikšić Theatre, who selected the play for production due to its status as a classic of modern world literature, confirms this idea. In a brief conversation on January 10, 2024, he explained that the play was not a narrative directly about the war, but rather a tale illustrating the breakdown of a value system. Jelić and his team also sought a director with a new, fresh conception and impetus, leading them to choose Serbian director Jug Đorđević. However, there were differing views on the adaptation. Jelić deemed it advantageous to contextualize the play by integrating elements associated with the officers of the Yugoslav People’s Army and issues characteristic of the region. Conversely, Đorđević believed that this approach might not be suitable for the production (Marojević 2021b, 5).

Đorđević asserted that, beyond the myriad of themes in the work, the play fundamentally explores a class conflict that is essentially a “cultural, religious conflict between warring parties” (Marojević 2021b, 5). He mentioned that he and his team were interested in exploring the

people behind the conflict, “the grey area between black and white”, where heroes, regardless of their actions, will repent (Marojević 2021b, 4–6). The play is based on the Serbian translation by Milica Drašković, which was then adapted for performance in Montenegro by the dramaturge Tijana Grumić. The dramaturge shared the view that the work possesses a universal meaning that resonates with today’s audience, eliminating the necessity to localize or update the play (Marojević 2021b, 5).

Yet, even though the production adheres to the source material, a notable aspect of the play’s adaptation was the introduction of angry women alongside the angry men. The dramaturge Grumić emphasized that “the class question is inseparable from women’s struggle and emancipation” (Marojević 2021b, 5). In this adaptation, as Grumić explained, they did not want the women “to be reduced to the consequences of male actions; they are not merely portrayed as fighting over men, causing division; instead, they show solidarity and share suffering, from which strength and revolution emerge” (Marojević 2021b, 5). In addition to this, the set designer Rondović made an effort to move the set away from realism, placing it in a Victorian apartment (Marojević 2021b, 5).

Audiences responded positively to the performance, expressing appreciation for the actors. The lead actor Jovan Krivokapić deemed it essential to communicate Jimmy’s discontent with society, emphasizing that the primary source of the protagonist’s anger lies in his insecurity and fear of emotional pain, encouraging him to make pre-emptive strikes against others to shield himself from potential harm (Marojević 2021b, 6). The actress Maja Stojanović went deep into her portrayal of Alison, highlighting the character’s lack of self-awareness (Marojević 2021b, 6). The play garnered acclaim, yet a notable challenge surfaced regarding its duration – a substantial 2 hours and 15 minutes, deemed excessive for the narrative at hand. Even though the play had 10 performances on both main and regional stages, the unresolved set-design issues impacted both logistics and audience engagement.

The most recent production by Nikšić Theatre is the contemporary Anglophone play *The Lungs* (2011) by Duncan Macmillan. It is a play that delves into the challenges faced by today’s younger generation, unfolding a distinctive narrative about childbirth, ecology, and the overarching theme of transience. The key factors influencing the selection of the play were its popularity, its suitability for the chosen format, and its relevance to contemporary themes. In a brief discussion in Nikšić Theatre on January 10, 2024, the art director Janko Jelić revealed that the play’s selection stemmed from discussions with the theatre’s actors. They sought a smaller format suitable for two or three actors, inspired by the success of the award-winning production of *The Lungs* in a co-production by Exit Theatre and City Theatre Sisak in Croatia. The conflict in the story appears when a modern, educated, environmentally conscious urban couple decides to have a child. As the narrative unfolds, the couple grapples with the complexities that arise when *he* suggests having a child, triggering *her* worries and a desperate need to overthink and rationalize.

All the uncertainties and dilemmas faced by young couples today as they plan their future resonate with Montenegrin audiences: Should they have children or focus on caring for the future of mankind? Is being good equal to caring about the environment? Are they environmentally conscious enough? Does the planet only need the genes of thoughtful,

educated, and caring people? How can they avoid the consequences for their children brought about by a flawed educational system or their parents' mistakes in upbringing, such as fitting children into certain gender roles or projecting their ambitions onto them? Various fears – such as the fear that having a child will signal the end of their own lives, the fear of not feeling love for the child, the fear of the effects of pregnancy on the human body, the fear of having a child with disabilities, and the fear of miscarriage – are just examples of the too many rationalizations that cause a lack of understanding between the couple. The key concern, however, is an existential question: how can anyone dare to bring a child into a world filled with serious problems such as overpopulation, lack of food and water, global warming, natural catastrophes, unpredictable climate change, political unrest, economic collapse, and an uncertain future. Although the existential crisis is overcome by the need for closeness when life happens, many dilemmas still remain. While written almost a decade before the COVID pandemic, the play's title, its references to suffocation, and a foreshadowing of a global catastrophe resulting in a population decrease – “what we need is the planet to fucking purge us, fucking drown us, burn us, cull everyone by about two-thirds” (Macmillan 2011, 13) – all allow for a new interpretation of the play.

The play received acclaim from critics, securing the 2013 Best New Play award at the Off West End Awards. According to the critic Lyn Gardner (2011), this “distinctive, off-kilter love story” is “brutally honest, funny, edgy, and current,” and is also “bravely written” and “startlingly structured”. The stage production in Nikšić was based on a translation by the Serbian theatre director Ivan Vuković. The stage adaptation adhered closely to the original text, with only minor changes. It premiered at Nikšić Theatre on November 30, 2023 under the direction of Andrej Nosov and got an excellent response from audiences, both in Nikšić and in the capital, Podgorica, mostly due to the play's contemporary themes being easily relatable, and the excellent performances. In a short span of time, the production has already had ten showings throughout Montenegro.

### 3.4 The City Theatre Podgorica

The City Theatre Podgorica itself, in line with its programming policy of staging internationally popular comedies aimed at meeting the demands of the audience (Tripković-Samardžić 2019, 226), has for many years included in its repertoire the contemporary Anglophone play *The Lying Kind* (*Lažljivci*), which received its 15th-anniversary performance in this theatre in March 2023. The adaptation, based on Anthony Neilson's play, was directed by Marko Manojlović.

As stated in the announcement of the City Theatre Podgorica, *The Lying Kind* (2002) is “one of the funniest, blackest and most shocking comedies ever seen” (Gradsko pozorište Podgorica n.d.). It was written by Scottish playwright Anthony Neilson, associated with In-Yer-Face Theatre.<sup>3</sup> The play, which portrays a world made up of a tangled web of lies and controversial

<sup>3</sup> Despite being initially linked to the *In-Yer-Face Theatre* (a term coined by Aleks Sierz and defined by Oxford Reference as “a new wave of British drama of the 1990s that was notable for its provocative uses of obscene language, nudity, violence, and taboo subject-matter” [n.d.], Neilson's later works, though still capable of startling audiences, harken back to the surreal and absurdist drama of an earlier tradition. See Bull (2011).

issues such as paedophilia, transvestitism, child abuse, dementia, and deaths of humans and animals, has been well received by Montenegrin audiences. An insight into the reception of the play in Montenegro in the first decades of the twenty-first century (Tripković-Samardžić 2019, 227) shows that the popularity of the play could be attributed to its current relevance. The plot, centred around a white lie concocted by police officers to shield elderly parents from their daughter's death, initiates "a whole chain of deceptions, provisional interpretations, grotesque situations, and errors" (Tripković-Samardžić 2019, 227). Despite the challenge of adapting British black humour for the Montenegrin context, the director Marko Manojlović utilized the black comedy genre to deal with such serious issues and present Neilson's scathing societal critique. The fact that this play continues to be performed on the stages of this theatre attests to the relevance of the themes it addresses.

## 4 Concluding Remarks

Theatrical responses to the challenges of the post-pandemic era in Montenegro affirm an ongoing trend toward the politicization of the theatre. The complex social, political, and cultural reality of Montenegro over the past three years, marked by two key traumas – grappling with the aftermath of the COVID pandemic and the polarization of society after the end of the decades-long rule of one political party – compounded by the global problems of the post-pandemic era, were reflected in Montenegrin theatre. Reflecting on Montenegrin theatre today and the significant gap between the exceptional talent of Montenegrin artists and the challenging working conditions they face stemming from structural and political issues, as well as the general politicization of the theatre, the director Paolo Magelli expressed hope that, despite the censorship that excludes theatre from the media and reduces it to its bare essentials, the theatre will strengthen its role as a place for enhancing "dialogue between the ensemble and the intelligence of the city where it operates," and become a space for "constant critique and free thinking" ("Revolucija je začeta u teatru" 2022, 6–7).

The nine contemporary Anglophone plays adapted for the stages of four Montenegrin theatres in this period have played a significant role in this dialogue. Given that the environment of transition is a constant feature of the theatre of the Balkans (Ljumović 2019, 141), it should not be surprising that these plays addressed the most burning issues of Montenegrin transitional society. Aimed at generating greater audience involvement, they managed to challenge the accepted notions on which this society is based.

Some of them, such as Lodato's *The Woman Who Amuses Herself* or McDonagh's *The Pillowman*, prompt us to reconsider our relationship to art in the era of cheap spectacle promoted by the traditional media in Montenegro, and redefine the role of art and its therapeutic function. The prevalence of female narratives, such as Stone's *Yerma*, not only reflects the contemporary struggles and dilemmas faced by women, but also echoes the specific, degraded position of women in the Montenegrin context. Others, such as Prebble's *The Effect* or Macmillan's *The Lungs*, reflect the uncertainties of young audiences and their existential despair in the face of the world's challenges and the Montenegrin crisis. Innovative elements in new readings of certain classics – such as the emphasis on the psychological perversions of the manipulated individual in Mamet's *Sexual Perversity in Chicago*, the focus on the angry women in Osborne's

*Look Back in Anger*, or the revolutionary undertones of Albee's *Who's Afraid of Virginia Woolf?* – provoke a revolt against liberal capitalism and its implications in current Montenegrin political reality. Sometimes, the only way to address the ills that pervade a corrupted society seems to be through black humour, as exemplified in the adaptation of Neilson's *The Lying Kind*.

The increased inclusion of contemporary Anglophone plays in the repertoires of four Montenegrin professional theatres, coupled with their popularity among theatre-goers in the country, serves as evidence of these theatres' efforts to contextualize issues relevant to their community, diversify the theatrical experience for audiences, and (re)establish the theatre as a place of debate. This is all the more apparent in light of the redefined role of a modern-day theatre as "a space of collective therapy, bringing forth symbolic reconciliation, (possibly) originating at the micro-level of imaginary community that comes into being for the duration of the performance between the performer and the audience" (Ljumović 2019, 142).

## References

- Bigsby, Christopher William Edgar. 2001. *Modern American Drama, 1945–2000*. Cambridge University Press.
- Bull, John. 2011. "Anthony Neilson." In *The Methuen Drama Guide to Contemporary British Playwrights*, edited by Sierz, Aleks, Martin Middeke, and Peter Paul Schnierer, 343–63. London: Methuen Drama.
- Crnogorsko narodno pozorište. n.d. "'Dokonda' izvedena na 26. Jugoslovenskom festivalu 'Bez prevoda' u Užicu." Accessed June 1, 2023. <https://cnp.me/dokonda-izvedena-na-26-jugoslovenskom-festivalu-bez-prevoda-u-uzicu/>.
- Čukić Šoškić, Iva. 2021a. "Šta se dogodilo Vićencu Perudi?" *Pozorište* 65 (42): 9–10. <https://niksickopozoriste.me/wp-content/uploads/2022/01/42-SA-KORICOM.pdf>.
- . 2021b. "The Effect: Pozersko pozorište." *Peripetija.me*, June 11, 2021. <https://peripetija.me/critic.194.critic.html>.
- . 2023. "Jerma: Drama braka bez djece." *Peripetija.me*, January 11, 2023. <https://peripetija.me/critic.224.critic.html>.
- Ćetković, Svetlana. 2022. "Pozorište se uvijek obraća savremenici." *Dan*, December 31, 2022.
- European Commission. 2023. "Commission Staff Working Document, Montenegro 2023 Report." November 8, 2023. [https://neighbourhood-enlargement.ec.europa.eu/system/files/2023-11/SWD\\_2023\\_694%20Montenegro%20report.pdf](https://neighbourhood-enlargement.ec.europa.eu/system/files/2023-11/SWD_2023_694%20Montenegro%20report.pdf).
- Gardner, Lyn. 2011. "Lungs – Review." *The Guardian*, October 10, 2011. <https://www.theguardian.com/stage/2011/oct/25/lungs-crucible-sheffield-review>.
- Gradsko pozorište Podgorica. n.d. "Lažljivci." Accessed January 13, 2024. <https://pgpozoriste.me/predstave/7>.
- Kiberd, Declan. 2006. "Reinventing England." In *A Companion to Modern British and Irish Drama 1880–2005*, edited by Mary Lockhurst, 22–35. Blackwell Publishing.
- Kontiđ, Jelena. 2021. "Individualni bijeg od društvenih konvencija." *Vijesti*, November 2, 2021. <https://www.vijesti.me/kultura/573636/individualni-bijeg-od-drustvenih-konvencija>.
- Koprivica, Dragan. 2021. "Rediteljski zlatni presjek scenskih parametara." *Pozorište* 65 (42): 7–8. <https://niksickopozoriste.me/wp-content/uploads/2022/01/42-SA-KORICOM.pdf>.
- Ljumović, Janko. 2019. "Contemporary theatre production in Montenegro and challenges of transition: Redefining programming policy." *Arc Academica* 7: 140–49.
- . 2021. "Karantinski eksperimenti." *Protagonist* 1 (September): 4–5. <https://zetskidom.me/wp-content/uploads/1884/01/Protagonist-broj-1.pdf>.

- . 2023. "Obnoviteljski duh CNP i izazovi savremenog pozorišta." In *Crnogorsko narodno pozorište 1953–2023*, 171–87. Podgorica: Crnogorsko narodno pozorište.
- Macmillan, Duncan. 2011. *Lungs*. London: Oberon Books.
- Marojević, Slavojka. 2020. "Pozorište u doba korone." *Pozorište* 64 (39–40): 21–22. <https://niksickopozoriste.me/wp-content/uploads/2022/01/Pozoriste-casopis-39-i-40.pdf>.
- . 2021a. "Priča o emotivnoj turbulenciji." *Pozorište* 65 (44): 13–14. <https://niksickopozoriste.me/casopis-pozoriste/>.
- . 2021b. "Gnjevom protiv samoće." *Pozorište* 65 (45): 4–6. <https://niksickopozoriste.me/wp-content/uploads/2022/01/Pozoriste-br-45.pdf>.
- . 2021c. "Čemu danas umjetnost služi." *Pozorište* 65 (42): 6. <https://niksickopozoriste.me/wp-content/uploads/2022/01/42-SA-KORICOM.pdf>.
- . 2021d. "Regionalna praižvedba." *Pozorište* 65 (43): 6. <https://niksickopozoriste.me/wp-content/uploads/2022/01/FINAL-43-SA-KORICOM.pdf>.
- Mrdenović, Maja. 2020. "Pozorište nedovršene tranzicije." *Peripetija.me*, June 29, 2020. <https://peripetija.me/.10.tnews.html>.
- Munjin, Bojan. 2022. "Treba se vratiti na početak." *Protagonist* 2 (3): 5.
- Nelević, Nataša. 2023. 2021. "Ko se boji Virdžinije Vulf: Ni milimetar iznad praga." *Peripetija.me*, November 7, 2021. <https://peripetija.me/critic.200.critic.html>.
- . "Jerma: Materinski zov našeg doba." *Peripetija.me*, January 9, 2023. <https://peripetija.me/critic.223.critic.html>.
- Oxford Reference. n.d. "In-Yer-Face Theatre." Accessed December 19, 2023. <https://www.oxfordreference.com/display/10.1093/oi/authority.20110803100010565>.
- Rakočević Zoran. 2021. "Niko više normalno ne radi one stvari." In *Čikaške perverzije* (official booklet). Cetinje: Kraljevsko pozorište Zetski dom.
- "Revolucija je začeta u teatru." 2022. *Protagonist* 2 (2): 6–7. <https://zetskidom.me/wp-content/uploads/1884/01/Protagonist-broj-2.pdf>.
- S. Vi. 2023. "Suočavanje djetinjstva i stvarnosti." *Pobjeda*. September 15, 2023. <https://www.pobjeda.me/clanak/suocavanje-djetinjstva-i-stvarnosti>.
- Stone, Simon. 2018. "With 'Yerma,' witness a woman unravel inside a terrarium." *Surface*, March 30, 2018. <https://www.surfacemag.com/articles/simon-stone-talks-yerma-set-design/>.
- Tripković-Samardžić, Vesna. 2019. "Contemporary Anglophone theater on Montenegrin stages from 2000 to 2015." *Coup de Théâtre: Crossing Borders: Anglophone Theater in Europe* 5: 219–35.
- Višnjic, Svetlana. 2023. "Osuda je omiljeni hobi ovog mentaliteta." *Pobjeda*, February 6, 2023.

## Metatheatre in Brian Friel's *The Loves of Cass McGuire*: The Semiotics of Make-Believe

### ABSTRACT

The purpose of the present paper is to explore the metatheatrical techniques in *The Loves of Cass McGuire* by Brian Friel. Those include the play-within-a-play, expressionism, as well as the self-reflexive techniques of Brecht's and Pirandello's metatheatre. In terms of structure, character delineation, language and setting, *Cass McGuire* draws attention to Friel's self-conscious devices. The play dramatizes the frustrations of Cass who deliberately chooses to withdraw to the realm of illusions and dreams, after being disappointed at the unpleasant realities she found at home. The study will also show how Friel managed to create a theatricalized balanced frame that draws the audience's attention to the theatrical devices and structural features of his play on the one hand, and to the reality beyond the theatre, or the world as "a projection of human consciousness", as Abel asserts, on the other.

**Keywords:** Brian Friel, metatheatre, semiotics in literature, Irish theatre, modern drama

### Metagledališkost v Frielovi drami *The Loves of Cass McGuire*: semiotika pretvarjanja

#### IZVLEČEK

V prispevku obravnavam metagledališke tehnike v drami *The Loves of Cass McGuire* Briana Friela, in sicer igro-v-igri in ekspresionizem, ter samorefleksivne tehnike Brechtovega in Pirandellovega metagledališča. Samozavedne strategije v tej drami so razvidne v dramski strukturi, zarisu likov, jeziku in prizoriščih. Drama tematizira frustracije junakinje Cass, ki se, razočarana nad neprijetno resničnostjo, ki jo najde doma, umakne v svet sanj in iluzij. V študiji tudi pokažem, kako je Frielu znotraj gledališkega medija uspelo ustvariti uravnotežen okvir, ki po eni strani pozornost občinstva usmerja na gledališka sredstva in strukturne lastnosti drame, po drugi strani pa tudi na resničnost onkraj teatra oziroma na svet kot Abelovo "projekcijo človeške zavesti".

**Ključne besede:** Brian Friel, metagledališče, semiotika v književnosti, irsko gledališče, moderna drama

# 1 Introduction

As early as 1940, the renowned Czech semiotician, aesthetician and theatre theoretician Jiří Veltruský declared: “All that is on the stage is a sign” (1940, 84). If semiotics is the study of signs, symbols and signification, or simply how meaning is created, then theatrical performances are at the heart of that science, as any theatrical presentation is a complicated act of semiosis. The work of the Russian formalists in the early decades of the twentieth century paved the road for the rise of semiotics in the field of literary criticism. Prominent among them were Viktor Shklovsky and Roman Jakobson, who drew the attention to the study of art as a science and laid great emphasis on the process of how art is created. They called for the concept of defamiliarization in art creation. In “Art as Technique”, Shklovsky states:

The purpose of art is to impart the sensation of things as they are perceived and not as they are known. The technique of art is to make an object “unfamiliar,” to make forms difficult, to increase the difficulty and length of perception because the process of perception is an aesthetic end in itself and must be prolonged. *Art is a way of experiencing the artfulness of an object; the object is not important.* (1998 [1917], 18)

The manipulation of semiotics in the analysis of theatrical forms was later developed by the Prague School in the 1930s and early 1940s. As Elam puts it:

In the context of the Prague School’s investigations into every kind of artistic and semiotic activity – from ordinary language to poetry, art, cinema and folk culture – attention was paid to all forms of theatre, including the ancient, the avant-garde and the Oriental, in a collective attempt to establish the principles of theatrical signification. (1980, 4)

Many scholars, language philosophers, and theoreticians have emphasized the value of the paradigmatic and syntagmatic relations of theatrical signs. Roland Barthes suggested in 1964 that “the nature of the theatrical sign, whether analogical, symbolic or conventional, the denotation and connotation of the message – all these fundamental problems of semiology are present in the theatre” (1972 [1964], 262). Susan Wittig went further by emphasizing the functional employment of semiotics in drama analysis:

Semiotics is a sturdy and serviceable theory of art, useful particularly as a means of understanding a multi-media art form like the drama and valuable not only as a way of talking about play itself but as a way into the teaching of dramatic theory in the classroom. (1974, 441)

Metatheatre recently gained momentum in the writings of many scholars who are interested in applying semiotics in the field of literary criticism, particularly in the appreciation of dramatic and theatrical forms (see, e.g., James 2020; Macrae 2019; Smith 2018). William Egginton observes that “there can be no theatre that is not already a metatheatre” (2003, 74). The term metatheatre or metadrama was first coined by Lionel Abel in his book *Metatheatre, A New View of Dramatic Form* (1963) and it was further delineated in a group of essays under the title *Tragedy and Metatheatre: Essays on Dramatic Form* compiled and edited by Martin Puchner (2003). The term refers to the artistic aspects or devices of a dramatic presentation

that draw the attention to its *artificiality*, or its being an art. That concept is thought by many to have been first inspired by the work of the Russian Formalists, especially Viktor Shklovsky (1917) and Boris Tomashevskij (see Rosenmeyer 2002; Elam 1980). Both called for the greater visibility of the devices used in a work of art, and that an author should attempt to “lay bare” his devices in order to draw the attention of the reader to them (Lemon and Reis 1965, 93). According to Abel, metatheatre or metadrama is used to describe a self-reflexive drama or performance that draws the attention of the reader/audience to the theatricalities and artistic devices or the *medium* that the author manipulates to convey his themes. For Abel, metatheatrical plays are

theatre pieces about life seen as already theatricalized. By this I mean that the persons appearing on the stage in these plays are there not simply because they were caught by the playwright in dramatic postures as a camera might catch them, but because they themselves knew they were dramatic long before the playwright took note of them. (2003, 134–35)

Thus, metatheatre with its emphasis on the theatricality of the dramatic text and performance contrasts sharply with the tenet of illusionism that lies at the centre of realism. For Andréa Pérez-Simon, metatheatre stands for “multiple self-referential, anti-illusionist devices in twentieth-century dramaturgies” (2011, 1). The formal features that Abel stipulates in metatheatre include the play-within-a-play form, characters who tend to be self-conscious about their own theatricality, a sequence of action that has the quality of a dream, a self-centred language that is marked by mediative metalinguistic functions, and open-ended closures. As will be shown in the following pages, all those, among other characteristics, are to be found in Friel’s *The Loves of Cass McGuire*.

## 2 Metatheatre and Its Relation to Comedy

Critics vary in their view of the alienating effects of metatheatre. While Richard Hornby believes that the Brechtian devices may lead the audience to “the most exquisite of aesthetic insights, that is ‘estrangement’ or ‘alienation’”, he claims that such an experience may create a sense of “unease” in an audience (1986, 32). However, for Stephen Purcell (2018), the same trait of theatricality creates rather a sense of “delight” in the audience. He proves this through connecting theatricality with Koestler’s concept of *bisociation* in his book *The Act of Creation*, where he argues that the theatre audience’s mind keeps oscillating between two matrices: the “Now and Here” and the “Then and There”. As Koestler puts it: “It is this precarious suspension of awareness between the two planes which facilitates the continuous flux of emotion from the Now and Here to the remoter worlds of the Then and There, and the cathartic effects resulting from it” (Koestler 1976, 306). Purcell further explains that the stage figure “is tangling and confusing the two matrices to such an extent that he or she sometimes becomes hard to locate in one or the other” (2018, 17). This is what Koestler means in his definition of jokes as “universes of discourse colliding, frames getting entangled, or contexts getting confused” (1976, 40). The device of oscillating between the two planes, the “Now and Here” and the “Then and There” pervades the behaviour of the main character in *The Loves of Cass McGuire*, as will be explained below. On the part of the audience this theatrical

oscillation arouses a sense of baffled delight. The metatheatrical experience for the audience is explained by Świontek as “metaphorically stepping back from engagement with the fictional properties of the art work into a cooler, more detached stance that perceives the process of its construction as art” (2006, 121).

## 2.1 *The Loves of Cass McGuire* as a Metatheatrical Play

*The Loves of Cass McGuire* was first performed at the Helen Hayes Theatre, New York on October, 6th, 1966. Ulf Dantanus regards *Cass* as “a sister play” to Friel’s first play, *Philadelphia Here I Come* (1964) and maintains that: “[t]ogether they make up an extensive statement on the themes of emigration, love, and attachment to home and family” (1988, 101). Abel states that metatheatrical plays present the world as “a projection of human consciousness” (1963, 60), and *The Loves of Cass McGuire* is designed to have a central consciousness through which the audience are allowed to view the events.

The play revolves around Cass who returns home to Ireland after 52 years of exile in America. She ran away to America when she was 18 due to poverty and a bad relationship. There, in a new land, she had an affair with a disabled American, Jeff Oslan, the owner of the restaurant where she worked, who did not find it necessary for them to get married. They lived together in his two-room apartment until his death. The action starts with Cass coming home after 52 years of exile, driven by the hope of warm love and compassion in her family. However, her sense of isolation and estrangement increases when she – being a heavy drinker – begins to cause trouble for her brother Harry, a well-off business man who enjoys respectability in his bourgeois, middle-class community. The play opens with her family deciding to put Cass – against her will – in Eden House, a care home for elderly people, which Cass insists on calling a “workhouse”. In this way the home to which Cass comes back turns out to be as emotionally sterile as the one she escaped 52 years ago. Defeated in her pursuit of love and self-fulfilment, Cass learns from the inhabitants of Eden House to escape from the undermining realities around her and to slip into the world of imagination, fantasies and romantic false memories. Cass decides to compose her own fictionalized narrative of the events of her life; she speaks to her fellow residents of her “ten-roomed apartment” in New York, and how Harry’s children used to send her letters “as regular as the clock”. Like most characters in Eden House, she performs her own rhapsody and so envisages the happiness that the real world has denied her.

In this play, Friel innovated a metadramatic form that reaches beyond the frames of common theatrical presentations. Dwelling on the play-within-a-play form, expressionism, as well as the self-reflexive techniques in Brecht and Pirandello’s meta-theatre, Friel crystalized a de-familiarized form of comedy that baffled the audience’s expectations. In this play, Friel demonstrates his theatrical talents and exceeds the familiar forms of structure, character delineation, language and setting, in order to convey referential messages that are meant to comment on the social and economic conditions of the Irish milieu. *The Loves of Cass McGuire*’s theatricality strikes the audience from the first scene. Starting with a scene in the manner of Pirandello’s *Six Characters in Search for an Author*, Cass’s direct address to the audience, her switching between the two planes (the Now and Here and the Then and There), the play within the play technique, the dream atmosphere that prevails throughout

the performance, and the metalinguistic discourse are among the prominent metatheatrical features of the work.

## 2.2 Theatricality Between the Expressionistic Structure and Pirandello's *Teatro Grottesco*

*The Loves of Cass McGuire* is composed of three acts, with a rhapsody ending each. The theatricality of the play is asserted from the first moment, with Cass rushing to the stage, objecting to the arrangement of scenes and taking over the performance. The play draws on Pirandello's absurd play, *Six Characters in Search of an Author* (1921), where six strange people interrupt a rehearsal and declare to the Director that they are "unfinished" characters searching for an author to finish their stories. Friel's play takes the form of a play-within-a-play, and the structure clearly makes use of expressionism, with the dream atmosphere dominating the action. The performance is devoted to dramatizing the inner thoughts and psychological dilemma of the protagonist through a series of flashbacks with overlapping and disconnected scenes. The action is not limited to the time Friel gives at the beginning of each act. The memory sequences utterly distort the chronological order of events. However, the development of action reveals a growing tension towards a climax in a confrontation scene when Pat faces Cass with her sordid realities:

Pat: A skivvy – that's what you were – written all over you! And a drunken aul skivvy living in sin with a dirty aul Yank that kicked you out in the end... Didn't your father do the same in Scotland...I know youse all right – tramps turned respectable." (Friel 1967, 44)

The anticlimax shows the broken Cass after losing her strong will to keep in contact with the present (or the audience). By the end of Act III, she has totally slipped into the realm of dreams and world of make-believe. As Pat expected, Cass goes downstage and sits in a winged chair, absorbed in a rhapsody about her wedding party, her happy life with Jeff in the ten-room apartment, the warm welcome she received when she and Jeff came home. The play ends with Cass arriving at a state of "calm satisfaction" – "Home at last. Gee, but it's a good thing to be home" (Friel 1967, 63). As Maxwell puts it, "The fluidity of time and setting, equally functional, corresponds to Cass's own vacillation as past and present threaten to coalesce into dream" (1973, 74). Thus, in a fine balance between theatricality and mimeses, Friel deftly blends expressionistic devices with Pirandello's *Teatro Grottesco*, to keep his audience ever fluctuating between the two planes: Here and Now, and There and Then.

## 3 Characters

### 3.1 Cass as the Author and Stage Director

The main character in Friel's play is responsible for the main part of theatricality in the whole performance. From the start she, as Abel demands, shows her awareness of being "dramatic long before the playwright took note of [her]" (2003, 134–35). Early in the opening moments of the play, Friel shatters the illusion of reality and allows Cass to directly address the audience and assume the role of the author and stage manager of the play. The play opens with a cozy

domestic scene in Harry McGuire's living room, two weeks before Christmas. Harry and his wife Alice are talking about the trouble Cass caused in the pub the night before while drunk. Suddenly, the illusion of reality is shattered with the arrival of Cass "who charges on stage (either from the wings or from the auditorium)". As Friel directs the scene, in her "raucous Irish-American voice" she objects to the way Harry chooses to arrange the events in a story about her life, and wants them to begin from the point she was "stuck in the gawddam workhouse" (Friel 1967, 12). She chases the McGuires away from the stage shouting:

Cass: What is this gawddam play called? *The Loves of Cass McGuire*. Who is Cass McGuire? Me! Me! And they'll [the audience] see what happens in the order I want them to see it; and there will be no going back into the past! (Friel 1967, 14)

Moving away, Harry says to Cass: "[...] you may think you can seal off your mind like this, but you can't. The past will keep coming back to you" (Friel 1967, 14). Now she owns the stage and decides to begin the story later on, at the point of her being put in the care home. She speaks directly to the audience: "They are her friends, her intimates. The other people on stage are interlopers" (Friel 1967, 12).

The fluidity of the play's setting adds to its metatheatrical atmosphere and grants Cass full control over the show. The curtain rises on a "spacious, high-ceiling room [...] which serves as the common-room in Eden House [...] and also as the living-room in the house of Harry McGuire" (Friel 1967, 7). When Cass charges onto the stage and interrupts the first scene at Harry's house, she immediately transforms the place into the home for the elderly: "Cass [looking around]: Yeah, this'll do for the workhouse. We have Swank windows, too, opening out on to a garden" (Friel 1967, 13). Quite as significant is her control of the lights and shadows that dramatize flashbacks. When Cass takes over the role of a stage manager, light is at hand to assist her: "[...] and we'll start off later in the story from here [light up bed-area]" (Friel 1967, 14). In this way, Cass carries the metatheatrical flair in Brian Friel's dramaturgy much further than any of his protagonists in the plays to come.

### 3.2 Characters as Dreamers

In expressionism, the characters are always presented as dreamers. Though the inhabitants of Eden House – unlike their expressionistic counterparts – are situated in a realistic locality, they create a world of make-believe for themselves. They indulge in reconstructing and transforming all the troublesome details in their lives into illusive images of happiness, success and fulfilment. With the help of those illusions, they are able to cope with their failures and acquire a sense of self-gratification.

Cass is presented as a tormented dreamer. She keeps shifting the audience's awareness between the two planes: the present reality and her memories. Like Gar O'Donnell who is divided into Public and Private in *Philadelphia*, she is divided between a past that she tries to escape from, and a present that she tries desperately to cling to. Cass does not want to acknowledge the fact that she has been rejected both by American society because of old age, and her Irish homeland because of her coarse manners. She yields instead to the teachings of her fellow residents in Eden House. Like them, she resorts to the make-believe trick through which she can cope with

the past. She manages, via self-deception, to reshuffle or reconstruct the unpleasant facts of the past into a romantic, pleasant memory that has nothing to do with reality.

With the exception of Pat, the rest of Eden House are also dreamers, and self-delusion dominates their lives to varying degrees. Ingram and Trilbe are the “rhapsodists” of Eden House who teach Cass the make-believe trick. They assist each other in living in and believing their illusions. Trilbe is a failed elocution teacher who lacks “the necessary qualifications” and was “consequently never recognized by the education department” (Friel 1967, 80). However, she refuses to acknowledge the fact. She puts on the mask of a distinguished elocutionist who is “adjudicating at a speech festival for junior schools next week” (Friel 1967, 19). However, one is conscious of “an insecurity behind the extravagant exterior” (Friel 1967, 18). Pat, another fellow resident, tells Cass that Trilbe has never married and that her father was a drunkard and that she, being a failed teacher, “kept running from one school to the next, hoping for a square meal” (Friel 1967, 28). In her rhapsody, Trilbe tells the others of her romantic love and marriage to a French prince, and how her father was “so proper and so stern”, and how they travelled around the world. She tells them of “the servant and the music and the wine and the travel and the poetry and his love for me and my love for him... all so real” (Friel 1967, 27).

Ingram is “a small, withered, testy nervous old man” (Friel 1967, 18). He was married to a dancer who abandoned him “two days after the wedding” (Friel 1967, 32) and eloped with a German Count in his yacht. This bitter experience affected his personality. As Friel directs him: “He is so frail and hesitant that he seldom finishes a sentence” (Friel 1967, 18). In his rhapsody, he tells us how his beloved “danced and danced and danced” for him, how they “kissed and loved and ran” and how she “one day, running before me, calling to me, she slipped” (Friel 1967, 41). Thus, in order to escape a painful past, each creates a myth of their own, believes it, and finally, lives it. Even Tessa, the young maid in Eden House – young as she is – catches the contagion of make-believe from the residents of the house. She tells them that her fiancé, a bricklayer, is a building contractor who will build “a bungalow” for her. Even the tangible facts about her wedding ring are transformed:

Tessa (by rote): It’s a solitaire diamond surrounded by a cluster of dazzling rubies and mounted on plat-ig-num and gold [Cass catches her hand and searches earnestly].

Cass: Where is the diamond?

Tessa: God, are you blind, too! There!

Cass: Oh yeah-yeah – so it is. Gee, that’s nice, sweetie. (Friel 1967, 63)

By the end of the play it is made clear that the residents of Eden House, including the maid, are role-players or masqueraders who are “locked into their own solipsistic world” (Andrews 1995, 96).

## 4 The Three Rhapsodies as Ritual

In the “Author’s Note” to the play, Friel states: “I consider the play to be a concerto in which Cass McGuire is the soloist” (Friel 1967, 8). The author ends each act with a rhapsody that does not belong to the realm of reality. In these each of the three rhapsodists – Trilbe,

Ingram and Cass – sings their own song following the same accompanying ceremonies or rituals. Each should begin the rhapsody by sitting and relaxing in the “winged chair” down stage; the music is to “fade in gently”, “slowly and with growing volume”. In the first two rhapsodies, Trilbe and Ingram join in a duet complementing each other’s recitation of their romantic illusions. As a newcomer, Cass’s reaction towards their strange way of speech is “naked astonishment” (Friel 1967, 25). Later on, she gradually starts to tolerate the dreamers’ rhapsodies, as she moves into their world of fantasy. Now she no longer resists memories as she did in Act I.

The language of the rhapsodists is flooded with lyricism and poetic imagery. It is also characterized by phonic repetitions that add to the musical rhythm of the rhapsody. At the end of each rhapsody, they quote one of Yeats’s romantic poems:

Trilbe: But I, being poor have only my dreams...

Ingram: Our truth.

Trilbe: ...I have spread my dreams under your feet.

Tread softly because you tread on my dreams.

Ingram: Our truth. (Friel 1967, 27)

The word “dreams” is significantly followed by “our truth” to suggest that the world of dreams is so absorbing for them that it has gained such a power as to be their truth. After the rhapsodies, each of the three rhapsodists is transformed and enters into a state of “calm satisfaction” and happiness. In Cass’s case, she is driven to the rhapsodic world by the persistent encouragement of Trilbe and Ingram.

Trilbe: Tell us.

Ingram: Tell us.

Trilbe: Tell us. (Friel 1967, 58)

At the beginning, Friel directs the scene as follows,

Cass suddenly bursts into tears and drops into the winged chair. There she cries and groans, covering her face with her hands while the spasm lasts. Then, emerging from it, she sits up straight almost with nobility and very slowly lets her head come to rest on the back of the chair. (Friel 1967, 58)

The fit of tears and the sudden agitation that befall Cass as soon as she starts her rhapsody are the remnants of an awareness of the world of reality that is still flickering in her mind. She is then immediately calmed by Trilbe’s encouragement to go on.

Trilbe: No, no, go on, golden hair and patient eyes. And you had two dimples...

Cass (Relaxing): Yeah – yeah two dimples... (Friel 1967, 59).

Wagner’s music for *Tristan and Isolde*, a love story ended by the deaths of the lovers, accompanies her rhapsody. As Patrick Burke comments: “Because [...] of the pathos attendant on our insight that Cass is a casualty of disappointment, lost love and ultimately semi-senility, the parallel between her situation and that of Wagner’s lovers is also made very ironic” (1997, 21).

## 5 Self-Conscious Language

### 5.1 Discourse and Meta-Discourse

A highly significant feature of discourse in metatheatres is that it deals with language as an object, or a theme to be discussed in the theatrical presentation, a feature that dominates another play by Friel, *Translations* (1980), and is also characteristic, to a lesser degree, of *The Loves of Cass McGuire*. As Elam maintains: “In the drama, the metalinguistic function often has the effect of foregrounding language as object or event by bringing it explicitly to the audience’s attention in its pragmatic, structural, stylistic or philosophical aspects” (Elam 1980, 96). The language used in Friel’s play can be regarded as having such a metacommunicative function. Moreover, much of the characters’ speech is not meant for communication with each other, or even to tell the audience their stories. It is rather meant for commenting on the characters’ own usage of language, covering up their psychological dilemma, and directly passing certain messages to the audience. Trilbe is a failed elocutionist whose fake and pompous language is a way of not facing her failure. Moreover, addressing the audience with direct speech, as Cass does, is regarded by Elam as “an extreme of linguistic self-consciousness” which serves “to ‘frame’ the very process of character-to-character or actor-to-audience verbal communication, and so becomes part of a broader metadramatic or metatheatrical superstructure” (1980, 96). Like most of Friel’s plays, the theatrical discourse in *The Loves of Cass McGuire* is sometimes meant to comment on the Irish milieu, or the “metatheatrical superstructure” of his plays. In this way, the characters’ speech turns out to be a mediative medium between the author and audience, a trait that characterizes “highly ideational drama”, as Elam points out, “of which Hamlet is the classical instance and the plays of Pirandello perhaps the best modern representatives” (1980, 95).

Umberto Eco says that “[s]emiology always seems inclined to affirm not that we speak the language but that we are spoken by the language” (1985, 590). In the play examined here, the discourse features of many of the characters are mainly used to highlight their psychological states, like Cass, Trilbe and Ingram, who “seldom finishes a sentence” (Friel 1967, 18). The “psychological preconditions of a language”, as Wittig puts it, are referred to by many linguists starting from Saussure’s langue-parole dichotomy, Levi-Strauss and Charles Morris, and later developed by many including the French semioticians Roland Barthes, Tzvetan Todorov, Claude Bremond, and A.J. Greimas (Wittig 1974, 445).

### 5.2 A Play of Voices

Being “a play of voices” (Andrews 1995, 96), *The Loves of Cass McGuire* displays a variety of forms which detach it from the everyday language of naturalistic drama and highlight its metatheatrical effect. Remarkable among these is the lyrical poetic language of the three rhapsodists. Language, Andrews indicates, “becomes preferable to direct action, to reality, to genuine engagement; it becomes a mask” (1995, 99). Even when the characters speak outside the rhapsodies, their conversations are characterized by shifts in topic and incoherence, a fact which illustrates estrangement and communication failure. The following dialogue is a pertinent example:

Trilbe: I think perhaps my favorite piece for children is a little poem ... It is called "Clickety-clack" ...

Ingram: When he was a young barrister my father published a pamphlet attacking capital punishment.

Trilbe: We puff over meadows and rivers and streams /Till we come puffing gaily to the land of our dreams.

Ingram: ... He was a stern man, but I think he was a just man.

Trilbe: We are having ham for tea.

Ingram: Snow is forecast. (Friel 1967, 55)

The lack of communication is primarily manifested by the long turns of speech that Friel allows Cass, "the soloist" of his concerto. Since she finds no emotional bridges with others, Cass speaks to the audience or to herself, a remarkable device that was used earlier in *Philadelphia* to suggest solitude and estrangement; the same device is to be used later on in *Molly Sweeney* (1994) and *Faith Healer* (1997). The play is also characterized by a variety of accents and idioms. Cass's vulgar language is contrasted with the ordinary middle-class language of Harry and Alice, the falsely pedantic speech of Trilbe and the hesitant, seldom finished sentences of Ingram. The dissonant modes of speech among the characters demonstrate the theme of non-communication and contribute to the comic quality of the play.

Silence is another element that Friel uses effectively to dramatize frustration and depression, particularly in the scene of Pat's departure. Cass, Trilbe and Ingram feel emotionally depressed and frustrated at their inability to find a way out of Eden House. This sense is exacerbated by Pat's success at leaving the place and getting another chance to live among the people outside. However, this is never expressed by words. After his departure, there are moments of uneasy silences intercut with some highly unnatural talk that cannot defeat the silence. "The silence now is total, the depression complete" and "They cannot sustain talk any longer. Silence flows in and fills the room" (Friel 1967, 55).

### 5.3 The Wagnerian Total Artwork and the Macro/Micro-Signs

Eco states: "Semiology shows us the universe of ideologies, arranged in codes and sub-codes, within the universe of signs, and these ideologies are reflected in our preconstituted way of using the language" (1985, 591). Throughout his dramatic career, Friel developed his own concept of the theatre craft to include all the aesthetic facilities of music, dance and ritual, together with the spoken word, to create a "universe of signs" in each play. In most of his plays, Friel adopts the Wagnerian concept of language that is fit for the theatrical experience. Wagner advocated a theory of the "total artwork" in music drama where "[n]ot one rich faculty of the separate arts will remain unused" (Wagner 1895, 190). In Wagner's notion of the ideal art, music, poetry and dance are combined to produce an artistic expression of some pre-rational time when those three elements served to "emotionalize the intellect" (Wagner 1895, 190). On the level of language, the most outstanding supra-realistic element in *The Loves of Cass McGuire* are the three rhapsodies of Trilbe, Ingram and Cass. Some critics (White 1999; Burke 1997; Andrews 1995; Rollins and Rollins 1990) referred to Richard Wagner's "total

artwork” as the background against which Friel formed the three rhapsodies in the play. The three rhapsodies are performed as a ritual in which music, words and setting are combined to present an allegorical frame of an absorbing illusion. These rhapsodies have nothing to do with the real world, and are the best linguistic form to dramatize the state of self-delusion and withdrawal into the world of fantasy. Rollins and Rollins compare the three rhapsodies in *The Loves of Cass McGuire* to selected passages from Wagner’s operas *Tannhauser*, *The Valkyrie*, and *Tristan and Isolde*. “Friel, like Wagner before him, is experimenting in this play with the reciprocal relationship between dramatic action and tonal harmonies” (1990, 24).

While Wagner’s concept of art emphasizes the total effect of the work of art as a “macro-sign”, theatre practitioners or even the audience deal with every device and sign on the stage as a deliberately directed “micro-sign” in which meaning is encapsulated. Petr Bogatyrev, formerly a member of the Russian formalist circle states in this context: “on the stage things that play the part of theatrical signs [...] acquire special features, qualities and attributes that they do not have in real life” (1976 [1938], 35–36). Elam also states that many theorists of the Prague Linguistic Circle, tended to “view the performance not as a single sign but as a network of semiotic units belonging to different cooperative systems” (1980, 5).

*The Loves of Cass McGuire* teems with indicative signs. Many stage props are used by Cass as a mask behind which she can hide her sordid reality, or to otherwise provide an escape from it. These include very small things like cigarettes, drinks and items of make-up, such as at the end of Act Two, when Cass is completely “deflated” by Pat’s cruel words to her; “Cass, angry, sobbing rushes into her room, takes the bottle from under the mattress and drinks [...] then she lights a cigarette and then makes up” (Friel 1967, 44). In the garden of Eden House, “a Cupid statue (illuminated) is frozen in an absurd and impossible contortion” (Friel 1967, 7). This is symbolic of the strange and “absurd” emotional state of isolation and lack of love in which the residents of the house live. Equally significant is the winged armchair that is situated down stage right “conspicuous in its isolation” (Friel 1967, 7). It is put down stage and “never used throughout the play except during the three rhapsodies” (Friel 1967, 7), as Friel writes in his directions. Giovanna Tallone sees it as “a parameter of escape in space and time” (2010, 60). It symbolizes a refuge into the world of dreams, where the inhabitants of Eden House “fly” high to find consolation for their grief. Before her rhapsody, Trilbe expresses her wish to sit in the winged chair. “Trilbe: I haven’t sat in it for three whole weeks, and now I wish to remember [...] The past and all the riches I have, and all that nourishes me” (Friel 1967, 26). As Kowzan puts it, all theatrical signs are “voluntary [...] Even if they have no communicative function in life, they necessarily acquire it on stage” (1968, 60).

## 6 The Actor/Audience Communication: “Declaring the Game”

Abel (1963) refers to many premediated and composed forms in many of Shakespeare’s and Ben Jonson’s plays as metatheatrical. Those include prologue, epilogue, induction, play within a play and asides. Such devices in which an actor is required to “step out of his role and acknowledge the presence of the public” (Elam 1980, 56), are meant to comment on the action, character or on the very theatrical experience. Though they appear to be “breaking the frame” as Elam puts it, they are “licensed means of *confirming* the frame by pointing out the pure facticity of the representation” (1980, 56).

In his book *Why Is That So Funny? A Practical Exploration of Physical Comedy*, John Wright describes metatheatre as a context within which the actor is “declaring the game” to the audience (2006, 45). He compares theatre practices to the rules of a game which both actor and spectator share, while knowing that there is “no illusion, and all your actions will be valued for what they are rather than for what they imply. When you declare the game, you play it so as to have an effect on the audience” (2006, 46). However, in none of his plays does Friel reach such a degree of transparency with his audience, and in the play under study he creates a balance between the mimetic and non-mimetic.

In the metatheatrical context of *The Loves of Cass McGuire*, the audience are endowed with a performative role and become part of the “game”. They symbolize reality as opposed to the world of dreams into which Cass gradually lapses. Cass addresses the audience as if they are among the characters of the play. Although she says to Harry “I live in the present, Harry boy: Right here and now!” (Friel 1967, 14), Cass frequently resorts to her memories, or the “there and then” plane. Throughout the play, Cass is torn by a bitter struggle between the past (or the painful memories that keep haunting her) and the present (or the reality that she tries to stick to and is represented by the audience).

The battle between the past and present is undecided for a time, as memory sequences invade her mind and converge with her direct speech with the audience. Near the end of Act II Cass realizes that her family no longer wants her among them. Now broken and “on the verge of tears” (Friel 1967, 39), she goes on talking to the audience of whose existence she wants to assure herself: “Are you still out there?” (Friel 1967, 40). Cass’s self-deception is marked by her awareness of the audience gradually beginning to fade and her getting more and more involved with Trilbe and Ingram in their rhapsodies. The tension of Cass’s struggle with the past comes to a height when Trilbe appears through the window beckoning her to join them. However, Cass desperately tries to stick to the audience, or the present, “They think they’re going to run me back into the past but by Gawd they’re not! [...] Where are you? Stick with me” (Friel 1967, 44).

Gradually, she begins to lose ground and finally lays down her arms. This is asserted when “she hesitates, takes a few steps towards the footlights, shades her eyes, searches the auditorium. She sees nobody”, “And I could’ov swore there were folks out there. [Shrugs] What the hell” (Friel 1967, 53). This is a signal that Cass has already lost contact with the world of reality. The play ends with the arrival of Mrs Butcher, a newcomer to Eden House who, ironically enough, speaks to nobody but the audience, whom she addresses “confidentially”. In this way, while distancing the audience from the characters through metatheatrical devices, Friel involves them in the performance and dwells on their self-consciousness to transmit his message. Patricia Waugh (1984, 63) states: “to be successfully decoded [...] experimental fiction of any variety requires an audience which is itself self-conscious about its linguistic practices.”

## 7 *The Loves of Cass Mc Guire* as an Irish Metatheatrical Piece

Friel’s identity and background and the play’s Irish setting are thus well served through the usage of European metatheatrical techniques. The deteriorating social conditions and political tension of Northern Ireland, his birthplace, are constant themes in almost all his entire oeuvre.

Derry, the city where Friel lived, was a minority Catholic community that suffered severely from political repression and economic depression since 1922, when the south of the island became the independent state of Éire/Ireland, while the north remained part of the United Kingdom. The separation of the north and south remained among Friel's recurrent themes, as expressed through the concepts of split identity, exile, emigration, dislocation and failure in communication and family disintegration.

Like *Philadelphia*, *The Loves of Cass Mc Guire* takes the crisis of a divided Ireland with its social and economic failure as a background to its central concern. Desolation, economic decay and sexual repression are the driving forces behind young Cass's decision to leave home and seek self-fulfilment in a new land, a situation that is identical to that of Gar O'Donnell. This is why critics see *The Loves of Cass McGuire* and *Philadelphia* as twin plays. As Ulf Dantanus explains: "Together they make up an extensive statement on the themes of emigration, love, and attachment to home and family" (1988, 101).

As Anthony Roche puts it, Friel's work "continues to speak to and interpret an Ireland undergoing dizzying social change and political upheaval" (1999, vii). The economic progress in Northern Ireland in the mid-1960s is also reflected in the successful career of Harry McGuire. John H. Whyte tells us in "Ireland, 1966–82" that the economy of Northern Ireland in 1966 had been improving, and the Ministry of Commerce "was energetically attracting new industries [...] Northern Ireland became one of the main centers in Europe for this booming industry" (1984, 342). When in exile, Cass believed that the ten dollars she used to send her family were of good use. When she returns home, however, she finds out that her money was left untouched and that her family did not really need it. She comes to realize sadly that her money has not bought her the love and compassion she expected to find on coming home. Thus Cass's social and economic dilemma, both before leaving Ireland and after coming home operates as a form of metatheatre, offering meaningful messages to the Irish audience. As Seamus Deane remarks significantly:

No Irish writer since the early days of this century has so sternly and courageously asserted the role of art in the public world without either yielding to that world's pressures or retreating into art's narcissistic alternatives. In the balance [Friel] has achieved between these forces, he has become an exemplary figure. (1996, 22)

## 8 Conclusion

Despite it being only Friel's second attempt at a play, *The Loves of Cass McGuire* introduced much of its author's "surefooted" (O'Toole 1988, 230) theatrical prowess, and already made use of a variety of methods and techniques that would also appear in his later works. These include expressionism, the *Teatro Grottesco* of Pirandello as well as the art of Wagner and Brecht's epic theatre. It validated Friel's experimental efforts and confirmed his grasp of many innovative devices, a fact that rendered it a "gem" (McGrath 1999, 71) among the author's early works.

The different techniques and dramatic devices that Friel manipulates in *Cass* are perfectly compatible with the make-believe theme of the play. The use of flashbacks and expressionistic memory sequences suits the dreamlike effect that he aimed to produce; the comic mood, and

three rhapsodies which reflect the influence of by Wagner's idea of the "total artwork", lend the play a lyrical ritualistic flair and help engage the audience and involve it in the dramatic performance. In those scenes Friel manipulates words, gestures, music, lights and shadows to signify the phantasy world to which Cass has moved. In a fine balance with that effect, Friel's debt to Brecht's epic theatre shows itself in making Cass play the role of the narrator, though she differs from Brecht's narrators in that she herself is involved in the action. Pirandello's impact is manifested in Cass's direct speech to the audience, her exclamation against the author's choice of the play's title, and her insistence on rearranging the events of the play. Such self-conscious devices contribute to the comic effect and confuse the audience by distorting the conventional concept of the invisibility of the spectator who is pushed by Friel, into not only "ponder(ing)" (Frey 2015, 106) the action, but also being part of the performance. We may conclude by agreeing with Rosenmeyer (2002, 99) that "a piece of metatheatre is by definition or concomitantly a variety of commentary, [...] both the play and the characters, and the playwright, and ultimately the audience, share in that critical pursuit".

## References

- Abel, Lionel. 1963. *Metatheatre: A New View of Dramatic Form*. New York: Hill and Wang.
- . 2003. *Tragedy and Metatheatre: Essays on Dramatic Form*. Edited by Martin Puchner. New York: Holmes & Meier.
- Andrews, Elmer. 1995. *The Art of Brian Friel*. London: Macmillan Press.
- Barthes, Roland. 1972 (1964). "Literature and signification." Translated by Richard Howard. In *Critical Essays*, 261–67. Evanston: Northwestern University Press.
- Bogatyrev, Pétr G. 1976 (1938). "Semiotics in the folk theatre." Translated by B. Kochis. In *Semiotics of Art: Prague School Contributions*, edited by Ladislav Matejka and Irwin R. Titunik, 33–49. Cambridge, Mass.: MIT Press.
- Burke, Patrick. 1997. "As if language no longer existed': Non-verbal theatricality in the plays of Friel." In *Brian Friel. A Casebook*, edited by William Kerwin, 13–22. New York: Garland Publishing.
- Dantanus, Ulf. 1988. *Brian Friel. A Study*. London: Faber.
- Deane, Seamus. 1996. "Introduction." In *Brian Friel: Plays I*. London: Faber and Faber. [First published as *Selected Plays* in 1984.]
- Eco, Umberto. 1985. "Articulations of the cinematic code." In *Movies and Methods: An Anthology*, edited by Bill Nichols, Vol. 1, 590–606. Berkeley: University of California Press.
- Egginton, William. 2003. *How the World Became a Stage*. Albany: State University of New York.
- Elam, Keir. 1980. *The Semiotics of Theatre and Drama*. London: Routledge, Taylor & Francis.
- Frey, Gaby. 2015. *Private Goes Public: Self-Narrativisation in Brian Friel's Plays*. Tübingen: Narr Francke Attempto Verlag.
- Friel, Brian. 1967. *The Loves of Cass McGuire*. London: Faber and Faber.
- Hornby, Richard. 1986. *Drama, Metadrama and Perception*. Lewisburg: Bucknell University Press.
- James, Allan. 2020. "Linguistic layering and the play semiotics of postdramatic (and dramatic) theatre: The case of Disco Pigs by Enda Walsh." *AAA: Arbeiten Aus Anglistik Und Amerikanistik* 45 (2): 109–38. <https://www.jstor.org/stable/26974200>.
- Koestler, Arthur. 1976. *The Act of Creation*. Great Britain: The Anchor Press.
- Kowzan, Tadeusz. 1968. "The sign in the theater: An introduction to the semiology of the art of the spectacle." *Diogenes* 61: 52–80.
- Lemon, Lee T., and Marion J. Reis, eds., trans. 1965. *Russian Formalist Criticism: Four Essays*. Lincoln: University of Nebraska Press.
- Macrae, Mitchell. 2019. "'[A]dore my topless villainy': Metatheatrical rivalry in John Marston's *Antonio's Revenge*." *Early Theatre* 22 (1): 93–118.

- Maxwell, Desmond E. S. 1973. *Brian Friel*. New Jersey: Associated University Press.
- McGrath, Francis C. 1999. *Brian Friel's (Post)Colonial Drama. Language, Illusion and Politics*. Syracuse: Syracuse University Press.
- O'Toole, Fintan. 1988. "Island of saints and silicon: Literature and social change in contemporary Ireland." In *Cultural Contexts and Literary Idioms in Contemporary Irish Culture*, edited by Michael Kenneally, 11–35. Gerrards Cross: Colin Smythe.
- Pérez-Simon, Andrés. 2011. "The concept of metatheatre: A functional approach." *TRANS – Revue de littérature générale et compare* 11. <https://doi.org/10.4000/trans.443>.
- Purcell, Stephen. 2018. "Are Shakespeare's plays always metatheatrical?" *Shakespeare Bulletin* 36 (1): 19–35. <https://doi.org/10.1353/shb.2018.0002>.
- Roche, Anthony. 1999. "Introduction: The worlds of Brian Friel." *Irish University Review* 29 (1): vii–x. <http://www.jstor.org/stable/25511522>.
- Rollins, Ron, and Nita Rollins. 1990. "'The Loves of Cass McGuire': Friel's Wagnerian Music Drama." *The Canadian Journal of Irish Studies* 16 (1): 24–32. <https://doi.org/10.2307/25512806>.
- Rosenmeyer, Thomas G. 2002. "'Metatheater': An essay on overload." *Arion: A Journal of Humanities and the Classics* 10 (2): 87–119. <http://www.jstor.org/stable/20163888>.
- Shklovsky, Viktor. 1998 (1917). "Art as 'technique.'" In *Literary Theory: An Anthology*, edited by Julie Rivkin and Michael Ryan, 17–23. Malden: Blackwell.
- Smith, Alexandra. 2018. "The transgressive capacity of the comic: A merry death as an embodiment of Nikolai Evreinov's vision of theatricality." *The Modern Language Review* 113 (3): 583–609.
- Świontek, Sławomir. 2006. "Excerpts from *Le Dialogue Dramatique et le Metathéâtre* by Sławomir Świontek." Translated by Jenn Stephenson. *Journal of Dramatic Theory and Criticism* 21 (1): 129–44.
- Tallone, Giovanna. 2010. "The old lady says 'no': The language of denial in Brian Friel's *The Loves of Cass McGuire*." *Ilha do Desterro: A Journal of English Language Literatures in English and Cultural Studies* 58: 53–76. <https://www.redalyc.org/articulo.oa?id=478348696003>.
- Veltruský, Jiří. 1940. "Man and object in the theater." In *A Prague School Reader on Esthetics, Literary Structure, and Style*, edited and translated by Paul L. Garvin, 83–91. Washington: Georgetown University Press.
- Wagner, Richard. 1895. *Richard Wagner's Prose Works*. Edited and translated by William Ashton Ellis. London: Kegan Paul, Trench, Trübner & Co.
- Waugh, Patricia. 1984. *Metafiction: The Theory and Practice of Self-Contained Fiction*. London: Routledge.
- White, Harry. 1999. "Brian Friel and the condition of music." *Irish University Review* 29 (1): 6–15.
- Whyte, John H. 1984. "Ireland, 1966–82." In *The Course of Irish History*, edited by T. W. Moody, and F. X. Martin, 335–50. Cork: The Mercier Press.
- Wittig, Susan. 1974. "Toward a semiotic theory of the drama." *Educational Theatre Journal* 26 (4): 441–54.
- Wright, John. 2006. *Why Is That So Funny?: A Practical Exploration of Physical Comedy*. London: Nick Hern Books.



**Martina Kastnerová**  
University of West Bohemia, Czechia

2024, Vol. 21 (1), 299-309(316)  
journals.uni-lj.si/elope  
<https://doi.org/10.4312/elope.21.1.299-309>  
UDC: 929Sidney P:929Johannes Crato

## Archiater Caesarius: Johannes Crato as Philip Sidney's Forgotten Mentor?

### ABSTRACT

Philip Sidney's European sojourn (1572–1575) proved to be a formative experience, one that shaped his political and literary development. Unsurprisingly, it has received much commentary. However, one rather neglected sphere of influence on Sidney's education that deserves fuller attention is the coterie of high-standing and learned figures based in Central Europe, many of whom Sidney either corresponded with or met in person. This paper will examine – using the example of the German humanist Johannes Crato von Krafftheim (1519–1585), personal physician to three Holy Roman Emperors, as a case study – how Sidney might have been inspired by the unique continental area of irenicism and the flourishing of new ways of understanding Man and Nature by means of medicine and botany.

**Keywords:** Philip Sidney, Johannes Crato, Elizabethan literature, Early Modern intellectual network, Sidney circle

## Archiater Caesarius: Johannes Crato kot pozabljeni mentor Philipa Sidneyja?

### IZVLEČEK

Obdobje, ki ga je Philip Sydney preživel v Evropi (1572–1575), je pomembno zaznamovalo njegov politični in literarni razvoj ter bilo prav zaradi tega deležno mnogih obravnav. Slabše raziskan pa ostaja vpliv učenjakov in veljakov, ki so delovali v deželah Srednje Evrope. Z mnogimi od njih si je Sidney dopisoval ali jih celo osebno srečal. V tem prispevku raziščem vpliv nemškega humanista Johannes Crata von Krafftheima (1519–1585), ki je bil osebni zdravnik treh rimskih cesarjev, na Sidneyjevo zanimanje za irenicizem in nove poglede na Človeka in Naravo, utemeljene na medicini in botaniki.

**Ključne besede:** Philip Sidney, Johannes Crato, elizabetinska književnost, zgodnje novoveška intelektualna mreža, Sidneyjev krog

*You have toasted me in French wine in L'Ecluse's letter: I for the moment will answer you here in Viennese wine, and will toast you when I get home in Hungarian. You are toasted also by my fellow drinkers, Masters Dadius, Lingel, Blotius, [and] Languet.*<sup>1</sup>

(Kuin 2012, 688)

## 1 Introduction: Philip Sidney's Intellectual Network in Central Europe<sup>2</sup>

Philip Sidney (1554–1586) is considered one of the most iconic figures of the Elizabethan era. Originally, though, his aspirations were not literary but rather political and diplomatic, as evidenced by his journey to the Continent (1572–1575)<sup>3</sup> and subsequent correspondence, which attests to Sidney's interest in the political/religious tensions of the time. His most important European mentor was Hubert Languet (1518–1581), a well-connected French diplomat and supporter of the Lutheran reformer Philip Melanchthon. Languet served as a kind of social facilitator for Sidney, introducing his young protégé to a number of prominent scholars and dignitaries (Kuin 2012, xlv; Kastnerová 2020, 8–9).

Sidney's "Grand Tour" was a formative experience, one that shaped his political and literary development and has received scholarly attention (Stillman 2008; Lockey 2021; Kuin 2021, 819–20). Nevertheless, one rather neglected sphere of influence on Sidney's education that deserves fuller attention is the coterie of high-standing and learned figures based in Central Europe, many of whom Sidney either corresponded with or met in person. Combing through the existing correspondence (of which only some 400 letters remain), the relevant archival sources, as well as his all-too-short oeuvre, important and perhaps overlooked details on the man and poet can be revealed.

Sidney's travels round Europe may have inspired the young poet to-be and pointed him towards the emerging discourse of modern medicine and experimental investigations of nature and humanity. Continental Europe was in this respect unique, as was the specific era during which Sidney visited many western and eastern parts of the Continent, especially those ruled by the Habsburg monarchy. During the period of Maximilian II's irenic policy, one found here an unprecedented number of important intellectuals interested and engaged in medicine and botany. In particular, this paper will examine Sidney's relationships with the German humanist Johannes Crato von Krafftheim (1519–1585), personal physician to three Holy Roman Emperors.

Based on the extant correspondence and archival sources, members of the Sidney-Languet milieu can be divided into three groups. All of these figures knew either Languet or Joachim

---

<sup>1</sup> "Salutasti me Gallico vino in litteris Clusianis, ego hic Viennense interim respondebo, et domi meae Hungarico te salutabo. Salutant te mei conpotores domini Domini Dadius, Lingelius, Blotius, Languetus" (Georg Purkicher to Sidney, Vienna, 6 June 1576).

<sup>2</sup> Research for this article was supported by the project/foundation SGS-2022-019: "Performativity and Narrativity in Art", University of West Bohemia, Pilsen. I also wish to thank Professor Robert E. Stillman and the library staff in Strahov Library Prague for their assistance. My thanks also go to Durham University and their Residential Research Library Fellowship, which made the completion of the article possible.

<sup>3</sup> In 1577, Sidney was sent back by the Queen on diplomatic missions. These were thus politico-diplomatic journeys rather than educational trips.

Camerarius and belonged to a generation of educated scholars who were all active in Vienna, Nuremberg and Wittenberg during Sidney's tours of the continent. The first, and most influential group, comprises five distinguished figures who met and/or directly corresponded with Sidney: Johannes Crato von Krafftheim, Carolus Clusius, Georg Purkircher, Thomas Jordan, Johannes Aicholz, and Thaddeus Hájek. Secondly, there is a German classical scholar Joachim Camerarius, who probably never met or corresponded with Sidney but, nonetheless, features indirectly in the correspondence. The personality of Joachim Camerarius Sr. is here of key importance due to his involvement in introducing Sidney to the network of Protestant intellectuals in the Habsburg monarchy (Stillman 2015). The third group is made up of important individuals who circulated within the milieu but do not feature in the correspondence, either directly or indirectly. They are the physician, historiographer and collector Johannes Sambucus (1531–1584), the physician and reformer (and Philipp Melancthon's son-in-law) Caspar Peucer (1525–1602) and the physician, botanist and cartographer Paul Fabricius (1519/29–1589).

It is quite clear that Sidney was especially interested in the Czech Lands and, more generally, the Habsburg Empire, where he visited local intellectual centres. He found there an inspiring approach to the administration of public affairs and a congenial multid denominational atmosphere. Roger Kuin (2021, 819–20) notes that this curiosity was likely driven by religious and political interests. Sidney may have gotten close to the Union of Brethren and observed the co-existence of Protestant and Catholic elements.

After leaving France, Sidney headed first to the French-speaking parts<sup>4</sup> of the Holy Roman Empire through Heidelberg, Frankfurt, and Strassburg to the Viennese court of Maximilian II. The emperor's irenicist policies seem to have influenced Sidney's thoughts on religion and politics both in terms of representing a form of tolerance in a positive sense of the word, and by acknowledging denominational fragmentation and disunity of Christendom in a critical sense (Lockey 2021; Louthan 2006). In Vienna, Sidney encountered various intellectual friends of his mentor Languet, including the abovementioned Hagecius, but also Johannes Crato (1519–1585) (the protagonist of this paper), and Michael Lingel (†1585), a leading representative of the Viennese medical faculty. Sidney's correspondence also tells us about his meeting with Jean de Vulcob (1535–1607), the enormously cultured French ambassador. These meetings must have been most inspiring – as attested by the fact that after visiting Hungary and Italy, Sidney came back to Vienna (Kuin 2021, 812–15).

## 2 Johannes Crato: Archiater Caesareus<sup>5</sup>

Johannes Crato was born and educated in Breslau in what is now present-day Wrocław. Interestingly, it was in Breslau that Crato's closest friend Andreas Dudith (1533–1589)

<sup>4</sup> Which is not surprising given that he did not speak German. It is also why we find among his contacts people connected with these areas, including Languet's friend Carolus Clusius / Charlese de L'Ecluse (1526–1609), who later accompanied Sidney during his trip to Hungary. At that occasion, they also passed through Bratislava, the birthplace of another of Languet's intellectual friends, Georg Purkircher (1530–1578).

<sup>5</sup> The Greek term used for court physician, synonymous with its Latin counterpart *principium medicus*, also features in the title of Crato's *Consiliorum, et epistolarum medicinalium, Ioh. Cratonis a Krafftheim, archiatri caesarei, et aliorum praestantissimorum medicorum...* 1593.

sought refuge after leaving Krakow following his public fall from grace. Ennobled by Maximilian II and taking the surname von Krafftheim, Crato represents a remarkable figure in that he served as court physician to three Holy Roman Emperors. But he is also important for his connections with Central European Protestantism, particularly Lutheranism and the Hussite movement, the Unity of the Brethren (*Unitas Fratrum*). According to Stillman, Crato distinguished himself as an impressive conciliator during the reigns of Maximilian II and Rudolf II and “devoted much of his life to persuading the emperors to grant religious freedom to the Bohemian Brethren” (Stillman 2021, 44).

Supported by Breslau city council and various patrons, Crato was given the opportunity of furthering his academic studies at the University of Wittenberg. There he studied theology under the tutelage of the German Lutheran reformer Philip Melancthon, who encouraged him to study classical languages as well as medicine. From here he moved to Padua to study medicine. In 1550, after his graduation, he returned to Breslau as municipal physician to repay the faith shown in him by his hometown. However, when suspicion of his Lutheran associations was confirmed following his conversion to the Protestant faith, he was formally removed from his post. Despite this setback, his career was revitalized upon his appointment as court physician to Ferdinand I in 1560. Crato would also go on to serve his successors, Maximilian II and Rudolf II. He was afforded particular favour during the reign of Maximilian, a period in which, as we have learned, a culture of irenicism prevailed at the Viennese court (Louthan 2006, 93–98).

Crato was a prominent figure within the Camerarius-Languet circle and enjoyed close friendships with Andreas Dudith and the French theologian Theodor Beza (1519–1605).<sup>6</sup> He was also an immensely and widely respected figure, with the Jesuits and Gnesio-Lutherans his only reputed detractors. That he retained his post as chief physician to three Holy Roman Emperors in spite of his Lutheranism is perhaps the most compelling evidence of his excellence as a medical practitioner. In October 1576, in what would prove a pioneering event, Crato performed the first autopsy ever to be documented in writing upon the body of Emperor Maximilian II himself (Kuin 2012, xxxix). Despite anatomy having been well-established as an integral part of medical study following the landmark publication of *De humani corporis fabrica* (*On the Fabric of the Human Body*) by Andreas Vesalius in 1543, the merits of dissection had been much disputed up until the time of Crato’s celebrated autopsy (Murphy 2019, 76–81).<sup>7</sup>

Crato was a regular correspondent with many of the leading humanists and intellectuals of the time, and regarded as an important facilitator in the elite socio-political network. He also owned a highly esteemed library collection, about which we unfortunately know little.<sup>8</sup>

<sup>6</sup> Various Czech noblemen of the Hussite movement were frequently accommodated at Beza’s residence in Geneva. One of these was the young student Karel Žerotín, who would later become secular leader of the group. His tenure coincided with Andreas Dudith’s arrival in Paskov. In one of his letters to Sidney, Languet refers to Žerotín in glowing terms (Kuin 2012, 912).

<sup>7</sup> This paper focuses particularly on parallels between Crato’s and Sidney’s spheres of interest, drawing on the relevant archive sources (especially Jagiellonian Library in Krakow, Wrocław Library, and Strahov Library in Prague).

<sup>8</sup> On the fate of Crato’s library, see Gillet (2015 [1860]). Parts of the collection are held at Nostitz Library in Prague and at Wrocław University Library.



FIGURE 1. Johannes Crato's signature *supralibros*.



FIGURE 2. 46th emblem by Johannes Sambucus.

In one of his publications from 1571, Crato's signature *supralibros* (a rare type of *ex-libris* adorning the cover of a book) depicts a figure (1) that appears to be either Samson or Hercules killing a lion.

Featuring the motto *Irae modereris et ori* ("May you moderate your anger and speech") (Šípek 2012, 14), the image resembles the 46th emblem of the *Emblemata* by Crato's eminent friend, the polymath and physician Johannes Sambucus. Entitled *Irae malagma philosophia* and also containing a fearsome lion, Sambucus' motif claims the greatest menace not to be an animal but anger, which can only be cured through philosophy (Sambucus and Iunius Hadrianus 1565, 52). Indeed, it is tempting to speculate that Crato's belief in championing temperance and rejecting gossip was directly related to his private experience of court politics.

The majority of treatises in Crato's collection were unsurprisingly devoted to medicine and historiography, with volumes on subjects ranging from nutrition to syphilis and featuring commentaries on Galen and Hippocrates. The collection also contained a number of remarkable publications, most notably Christopher Plantin's polyglot bible, the *Biblia Regia*. Crato was also mentor and tutor to many young students, who were able to profit from his considerable resources and seasoned advice (Šípek 2012, 24–25).

The most notable of Crato's own publications are his edition of Paul Eber's *Calendarium Historicum* and a dedication letter composed for an edition of Dubravius' *Historia Bohemica* in 1575.



FIGURE 3. Eber's insignia in Paul Eber's *Calendarium Historicum*.



FIGURE 4. Joachim Camerarius' emblem *Si Serenus Illuxerit*.

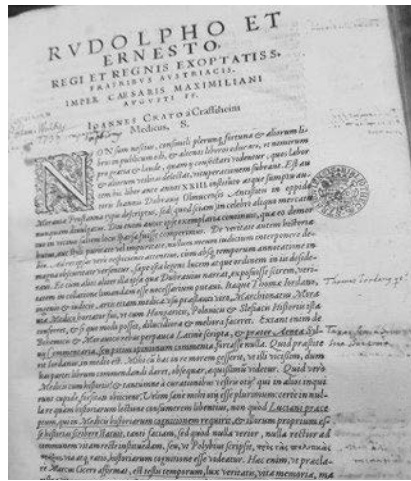


FIGURE 5. Crato's introductory letter of dedication.

Also based in Wittenberg, Paul Eber (1511–1569) was a German Lutheran theologian and reformer. Featured in Crato's edition of Eber's calendar published in 1571 is a verse from Psalm 119: "Thy word is a lamp unto my feet and a light unto my paths" (119:105, KJ21).<sup>9</sup> Crato probably understood this image of enlightenment very well, symbolic of a thematic trend in the output of many of the philosophers and intellectuals of the time. One such example is a rewording of Jacob 1.17 by Crato's friend, Joachim Camerarius. Featuring an emblem entitled *Si serenus illuxerit* (Camerarius 1677), it reads: "Every good gift and every perfect gift is from above, coming down from the Father of lights, with whom can be no variation nor turning shadow" (Jacob 1:17, B21).<sup>10</sup>

Published in Basel in 1575, Peter Perna's edition of Dubravius' *Historia Bohemica* contains Crato's introductory letter of dedication, which is notable for its addressing not only the ruler of the time, Maximilian II, but also his sons: his soon-to-be-successor Rudolf and Ernest (1533–1595), latterly governor of the Netherlands.

Written in 1574 in Vienna, the history was published, perhaps strategically, one year later, coming just a year before Maximilian's death. Significantly, Crato delivered an oration at Maximilian's funeral that would receive wide praise, particularly from his ally Sambucus (Crato 1577). Crato's *epistola dedicatoria* summarizes the important milestones in Dubravius' narration of Czech history. Quoting Cicero, he commends the merits of historiography and its enduring ability to bear truthful witness to memory and times past (Dubravius 1575, 2).<sup>11</sup> Crato concludes his dedication by accentuating Maximilian's virtues of foresight and tolerance, but particularly his dedication to avoiding conflict while simultaneously upholding

<sup>9</sup> "Lucerna pedibus meis verbum tuum: et lumen semitis meis."

<sup>10</sup> In Camerarius' Latin version: "Omne munus bonum et omne donum perfectum descendit desuper a Patre luminum."

<sup>11</sup> "[...] est testis temporum, lux veritatis, vita memoriae, magistra vitae, nuncia veritatis."

a passionate and brave defence of the state (Dubravius 1575, 5). Considering his close relationship to Maximilian, Crato's candid testimony here rings strikingly true. Indeed, his advocacy of Maximilian's diplomacy perhaps speaks of Crato's intuitive awareness that the age of irenicism was at an end.

Crato also published treatises on Lutheran and Protestant theology, medicine and astronomy, notably *Elementa Doctrinae De Circulis Coelestibus* (1551) by Caspar Peucer, a prominent member of the Wittenberg-Melanchthonian circle. He also compiled an informative and popular medical guide for households entitled *Euporista Cratoniana*, posthumously published in 1630 as a German translation under the title *Euporista Cratoniana, Oder Hauss Artzneyer*. The guide is evidence of the contemporary fashion for blending medicine with botany. Covering a diverse range of topics, it explores the medicinal uses of opium and provides advice on treating everything from scorpion and black-widow stings and burn marks to moles, skin peeling and dermatological infections. It also includes an extensive chapter on the induction of labour, delivery and other general aspects of childbirth (Crato 1630).

## 2.1 Sidney's Companions: Languet and Crato

Given Crato's exalted pedigree, Languet was naturally keen to introduce him to his young protégé. By the time of Sidney's move to Central Europe, Crato was a well-established and active member of the Languet-Camerarius circle, which comprised numerous botanists and physicians. We know from the correspondence that Sidney was put in contact with a number of these figures, which did much to inspire his interest in medicine and botany. Unfortunately, the only evidence of any direct communication between Crato and Sidney is one letter written by Crato in Latin. However, Crato is mentioned in three letters written by Hubert Languet to Sidney. The first two date to 1574, coinciding with Sidney's tour of the European mainland, and the third, written in 1577, dovetails with Sidney's brief diplomatic mission to Prague, where he was rumoured to have secretly met Edmund Campion.<sup>12</sup>

The first letter that mentions Crato, written by Hubert Languet in Vienna on 18 March 1574, concerns Sidney's intention to travel to Poland for the anticipated coronation. It is clear from the letter that Languet had initially planned on accompanying Sidney in spite of his diminishing state of health, being evidently more concerned with the welfare of his protégé than his own. But what is interesting is that Crato seems to have been familiar with Sidney's intention (and Languet's plans to accompany Sidney to Poland), a matter Languet insists he had kept a great secret (Kuin 2012, 140–42).<sup>13</sup> Languet's letter reveals, however, that Crato had written a message to Sidney disclosing his knowledge of Sidney's plans (Kuin 2012, 141, no. 2). The surprise felt by Languet, who considered himself Sidney's closest friend and mentor, is particularly evident in the letter: "I am surprised that Crato wrote what you say he did, as I mentioned that business here neither to him nor to anyone else" (Kuin

<sup>12</sup> Edmund Campion (1540–1581) was an Oxford scholar who famously rejected Anglicanism to become a Jesuit. His meeting with Sidney has generated much debate, with some interpreting the event as an indication of Sidney's religious tolerance and even, more speculatively, of his secret Catholicism.

<sup>13</sup> Hubert Languet to Sidney, Vienna, 18 March 1574.

2012, 140–42).<sup>14</sup> One can venture two possibilities for Crato having known. One is that Sidney himself informed Crato, which makes sense given that Crato was a trusted and close friend. Moreover, being of Polish-Silesian origin, Crato was ideally suited to offer advice on his upcoming trip. The other possibility is that the information found its way to Crato via a member of the imperial court, potentially Clusius or Purkicher. Whatever the truth of the matter, Sidney is more than likely to have found Crato's knowledge of the political scene in Poland and his acquaintance with the important figures of the region invaluable.

After recovering from an ailment that laid him low in September of 1574, Sidney eventually made the trip to Poland at some point between late October and early November of the same year. He is understood to have been accompanied by a Polish nobleman, Marcin Lezniowski, due to Languet's acceptance of a diplomatic post in service of Augustus, Elector of Saxony (Kuin 2012, 334). That Sidney's first stop was Breslau points to Crato's influence. Continuing to Krakow, the hub of Polish political, cultural and intellectual life, he met Andreas Dudith (Osborn 1972, 54; Kuin 2012, 208).

The second letter, written a month later on 23rd April 1574, refers to Crato as "a physician and old friend" (Kuin 2012, 173). The content of the letter mostly concerns Crato's connection with the polymath and reformer Jakob Monau. The letter is otherwise unexceptional save for the exchanges between Languet and Sidney on the tardiness of their replies and their mutual exasperation with the delivery of private letters (Kuin 2012, 173–77). The last letter referring to Crato comes more than three years later. Written in Frankfurt on 14th June 1577, Languet informs Sidney of a letter he has received from Crato. In it, the physician relays his warm greetings and expresses his regret at having missed the poet's most recent visit (Kuin 2012, 747–51). Although a snapshot of the relationship between Crato and Sidney, it nonetheless provides important evidence of the time Sidney spent in Central Europe, a period that coincided with Edmund Campion's visit to Prague. Secondly, it testifies to Crato's precise knowledge of Sidney's whereabouts in the region.

If we were to ask whether familiarity with Central European intellectuals influenced Sidney's work, we must, aside from the abovementioned historical circumstances, rely on a search of intertextual parallels. Let us therefore cite some examples.

In one of Johannes Sambucus's emblems, the *Physica & Ethica*, the investigation of nature and humanity is represented by the very figure of Johann Aicholz (1520–1588), a botanist and physician active in Vienna and Prague, and a friend of Paul Fabricius (1519/29–1589) and Carolus Clusius (1526–1609). The investigation of nature trains our minds' ability to recognize hidden things – nature does not easily give up its secrets. Characteristically, we find here an emphasis on *praxis*, that is, the application of theoretical knowledge to everyday life, to the real world. Sambucus supports this position with reference to Socrates: "but what good would be the use of ancient philosophy without suitable examples?" (Sambucus 1565).<sup>15</sup>

<sup>14</sup> "... Miror Cratonem scripsisse, quae dicis, cum nec apud ipsum, nec apud quenquam alium ejus rei mentionem hic fecerim."

<sup>15</sup> In Sambucus's Latin text: 'Huius alumna sed est praxis, mandata capessens / Notitiae, ac licitis usibus apta nitet. / Socratis haec placuit studio, quid enim sine rectis / Profuit exemplis Philosophia vetus?'; (<https://www.wdl.org/en/item/14211/view/1/7/>, accessed August 24, 2021, 150; English translation at <https://www.emblems.arts.gla.ac.uk/>

Sidney's *Defence of Poetry* is characterized by such close connection between a theoretical exposition and illustrative examples. The introduction starts with a depiction of Sidney's meeting with Giovanni Pietro Pugliano at the Viennese court, where the latter was so eloquent about his riding skills that he nearly talked Sidney into wishing for a horse. This eloquence and convincing effect were due to the syncretism of theory, or rather contemplation, illustrative examples, and great love for the subject of the discussion. This is something Sidney views as an example that should be followed in his own method of presentation (Sidney 2008, 212). It is interesting to note that Joachim Camerarius the Elder (1500–1574), father of the Camerarii and a close friend of Sidney's mentor Hubert Languet (1518–1581), was seriously interested in the art of riding, and for this reason translated and published Xenophanes' work on this subject (McMahon 1947, 88).<sup>16</sup>

It is clear that applicability to everyday life is of key importance for Sidney. When, with the aim to demonstrate its primacy, he compares poetry and its potential with that of philosophy and history, we can clearly hear a reverberation of his inspiration: in the Aristotelian spirit, he depicts poetry as the perfect counterpoint to the obscurity of philosophy and descriptive superficiality of historiography. He notes that “the poet is the food for the tenderest stomachs, the poet is indeed the right popular philosopher”, and that poetry “is more philosophical and more studiously serious than history”. In its moral impact and universality, poetry surpasses history. In this, it is comparable to philosophy but surpasses it in evoking pleasure and emotions (Sidney 2008, 232, lines 467, 477).

In Sidney's works we also find in many different guises a fascination or sympathy with a cyclic conception of time, inspired by continental emblematisers like above-mentioned Sambucus or Joachim Camerarius the Younger (1500–1574), where coming to be and decay are just part of the order of nature, and acceptance of this order is what brings peace. Several examples can also be found in Sidney's *Old Arcadia*. For instance, in the fourth eclogue Agelastus laments the vanity of human life, but in these verses Sidney also offers consolation by pointing to the restoration of life as such: in time, youth and old age always follow one another, but this cycle never ends with just one of these phases. A person cannot shed their old skin like a reptile; everyone ultimately heads towards death. Agelastus mourns the injustice of nature which permits the death of a child, but a closer inspection shows that the rhetorical nature and exaggeration in this speech are an instrument of subtle irony: dwelling on our individual pain does not bring peace. Peace of mind comes with accepting the order of nature to which we, as “nature's works”, are subjected (Sidney 2008, 131, lines 79–84).

On this note, emblem XLII by Camerarius – *Natura dictante feror* (I fly where nature bids me) – captures the ideal of Sidney's view of life, that is, the acceptance of the natural order of nature, *physis*, which should be investigated not only by contemplation: one should deploy the nascent Early Modern approaches to all of nature's beings, which are not only spiritual but also material entities. Such an investigation offers the option of being led by knowledge of one's own nature.<sup>17</sup>

---

french/emblem.php?id=FSAb106, accessed July 29, 2022).

<sup>16</sup> In hoc libello haec insunt. De tractandis equis sive... Tubingae Sueuorum (1539).

<sup>17</sup> More to Sidney's interest in continental emblem books in Kastnerová (2023).

### 3 Conclusion

The correspondence covered in this study illustrates that Crato was an influential and inspiring companion of the young Sidney, accompanying him in exploration of continental irenicist politics and the ideas of eminent local physicians and botanists like Clusius, Purkircher, or Aicholz. Considering his distinguished service at the imperial court, one can arguably claim Crato to be an equal contemporary of Languet in both age and social standing. The first letter not only highlights Languet's awareness of Crato's influence, but also his desire to maintain an exclusive relationship with his young English charge. The other letters offer an insight into the respect and admiration he commanded within the Languet-Camerarius circle, a group that was pivotal in paving the way for Sidney's integration into European intellectual life. But what is perhaps most revealing is the intimacy of Crato's friendship with Sidney, a closeness it would seem that went beyond mere words.

In summary, Crato's formative influence on Sidney's intellectual development is twofold. Firstly, he piqued Sidney's interest in the political affairs of Poland and the other states of Central Europe (to which Sidney also travelled). Secondly, by placing Sidney in direct contact with leading figures such as Clusius and Purkircher, he could inspire in the poet an interest in medicine and botany.<sup>18</sup>

### References

#### *Manuscripts and Old Prints*

- Camerarius, Joachim. 1677. *Joachimi Camerarii Symbolorum et emblematum centuriae quatuor: quarum prima Stirpium, secunda Animalium quadrupedum, tertia Volatilium et insectium, quarta Aquatilium et reptilium; rariores proprietates historias ac sententias memorabiles non paucas breviter exponit*. Mogvntiae: sumpt. Ludovici Bourgeat. Jagiellonian Library, Krakow, sign. BJ St. Dr. Neolat. 128.
- Crato, Johannes. 1593. *Consiliorum et epistolarum medicinalium Ioh. Cratonis a Kraftheim, archiatri caesarei, et aliorum praestantissimorum medicorum...* Frankfurt.
- . 1630. *Evporista Cratoniana Oder Hauß Artzneyer*. Wroclaw: bey Abraham Lambergs Seel. Erben. Wroclaw University Library, sign. 411760.
- Crato von Crafftheim, Johannes. 1577. *Oratio funebris de Divo Maxaemiliano II. Imperatore Caesare Augusto*. Vratislaviae: in officina typographica Crispini Scharffenbergii. Moravian Library, Brno, sign. ST1-0023.337.
- Dubravius, Johannes – Jordanus, Thomas. 1575. *Historia Boiémica. A Thoma Jordano, medico annotationibus ornata*. Basileae: apud Petrum Pernam. Strahov Library, Prague, sign. AO I 12/a.
- Eber, Paulus. 1571. *Calendarium historicum*. Vitebergae: Johannes Crato. Strahov Library, Prague, sign. EO XIII 2.
- Sambucus Johannes – Iunius, Hadrianus. 1565. *Emblemata*. Antverpiae: apud Christophorum Plantinum. National Széchényi Library, Budapest, sign. Ant. 7101.

#### *Correspondence and Belles-lettres*

- Kuin, Roger, ed. 2012. *The Correspondence of Sir Philip Sidney. Volume I, II*. Oxford: Oxford University Press.
- Sidney, Philip. 2008. *Sir Philip Sidney: The Major Works*, ed. Katherine Duncan-Jones. Oxford: Oxford University Press.

<sup>18</sup> Professor Daniel T. Lochman's (Texas State University) examination of Sidney's works, taking into consideration the poet's European travels and first-hand acquaintance with continental physicians, indicates that Sidney was influenced especially by paracelsianism (maybe also thanks to Thomas Moffet (1553–1604)). He could have been inspired by some thoughts on imagination, as we could see in his conception of nature and poetry (creation) in his *Defence of Poesy*. Thus, there are other directions for the future research of the topic of this study.

**Holy Writ**

Jacob 1:17 (B21)

Psalm 119: 105 (KJ21)

**Other Books and Articles**

- Gillet, Johann F. A. 2015 (1860). *Crato von Crafftheim und seine Freunde: Ein Beitrag zur Kirchengeschichte 1860*. Delhi: Facsimile Publisher.
- Kastnerová, Martina. 2020. "Johannes Sambucus at the imperial court and his ties to the Sidney circle." *Sidney Journal* 38: 1–28.
- . 2023. "Learned images of Nature: Physica, Philip Sidney and the Camerarius Brothers." *Sidney Journal* 41 (1–2): 121–33.
- Kuin, Roger. 2021. "Philip Sidney's travels in the Holy Roman Empire." *Renaissance Quarterly* 74 (3): 802–28. <https://doi.org/10.1017/rqx.2021.101>.
- Lockey, Brian E. 2021. "Philip Sidney and Edmund Spenser on transnational governance and the future of Christendom." *Renaissance Quarterly* 74 (2): 369–411. <https://doi.org/10.1017/rqx.2021.1>
- Louthan, Howard. 2006. *The Quest for Compromise. Peacemakers in Counter-Reformation Vienna*. Cambridge: Cambridge University Press.
- McMahon, Philip. 1947. "Sir Philip Sidney's Letter to the Camerarii." *PMLA* 62 (1): 83–95. <https://doi.org/10.1632/459194>.
- Murphy, Hannah. 2019. *A New Order of Medicine: The Rise of Physicians in Reformation Nuremberg*. Pittsburgh: University of Pittsburgh Press.
- Osborn, James M. 1972. *Young Philip Sidney, 1572–1577*. New Have: Yale University Press.
- Stillman, Robert E. 2008. *Philip Sidney and the Poetics of Renaissance Cosmopolitanism*. Aldershot: Ashgate.
- . 2015. "Philip Sidney, Thomas More, and table talk: Texts/contexts." *English Literary Renaissance* 45 (3): 323–50. <https://www.jstor.org/stable/48634682>.
- . 2021. *Christian Identity, Piety, and Politics in Early Modern England (ReFormations: Medieval and Early Modern)*. Notre Dame: University of Notre Dame Press.
- Šípek, Richard. 2012. "Knihovny raně novověkých lékařů ve fondu Nostické knihovny v Praze." *Acta musei nationalis Pragae* 57 (1–2): 24–25.

**Figures**

- Figure 1. Johannes Crato's signature *supralibros*. Accessed August 18, 2020. <http://files.provenio.net/PROVENIO/NOSTICKA/CRATO/IONB.jpg>. [Public domain.]
- Figure 2. 46th emblem by Johannes Sambucus. Emblema XLVI. Johannes Sambucus and Iunius Hadrianus, *Emblemata* (1564), 52. Accessed August 24, 2021. <https://www.wdl.org/en/item/14211/view/1/7>. [Public domain.]
- Figure 3. Eber's insignia in Paul Eber's *Calendarium Historicum* (Vitebergae: Johannes Crato, 1571); Strahov Library, Prague, EO XIII 2.
- Figure 4. Joachim Camerarius' emblem *Si Serenus Illuxerit*. Accessed June 17, 2022. <http://distichalatina.blogspot.com/2012/01/si-serenus-illuxerit.html>. [Public domain.]
- Figure 5. Crato's introductory letter of dedication. Peter Perna's edition of Dubravius' *Historia Bohemica* in Johannes Dubravius and Thomas Jordanus, *Historia Boiémica. A Thoma Jordano, medico annotationibus ornata* (Basileae: apud Petrum Pernam, 1575); Strahov Library, Prague, AO I 12/a.

All figures (except for that from public domain) by kind permission of Strahov Library in Prague.





# List of Contributors

# LIST OF CONTRIBUTORS

**Mahasen Badra**

Istanbul Gelisim University, Turkey  
mmmbmahasen@gelisim.edu.tr

**Sándor Czeglédi**

University of Pannonia, Hungary  
czegledi.sandor@htk.uni-pannon.hu

**Ifeta Čirić-Fazlija**

University of Sarajevo, Bosnia and Herzegovina  
ifeta.ciric@ff.unsa.ba

**Claudio Díaz**

University of Concepción, Chile  
claudiodiaz@udec.cl

**Nikola Dobrić**

University of Klagenfurt, Austria  
nikola.dobric@aau.at

**Katja Dobrić Basanež**

Juraj Dobrila University of Pula, Croatia  
kdbori@unipu.hr

**Paulo Etchegaray**

Autonomous University of Chile, Chile  
paulo.etchegaray@uautonoma.cl

**Michelle Gadpaille**

University of Maribor, Slovenia  
michelle.gadpaille@um.si

**Aleksandr Gaisov**

University of Primorska, Slovenia  
aleksgaisov@protonmail.com

**Martina Kastnerová**

University of West Bohemia, Czechia  
kastnerm@ff.zcu.cz

**Ivana Krsmanović**

University of Kragujevac, Serbia  
krsmanovici@gmail.com

**Bernarda Leva**

University of Maribor, Slovenia  
bernarda.leva@um.si

**Silvana Orel Kos**

University of Ljubljana, Slovenia  
silvana.orelcos@ff.uni-lj.si

**Mabel Ortiz**

Catholic University of the Most Holy Conception, Chile  
mortiz@ucsc.cl

**Vesna Petrović**

University of Kragujevac, Serbia  
vesna.petrovic@vstss.com

**Katja Plemenitaš**

University of Maribor, Slovenia  
katja.plemenitas@um.si

**Nastja Prajnc Kacijan**

University of Maribor, Slovenia  
nastja.prajnc1@um.si

**Marcela Quintana**

Arturo Prat University, Chile  
maquinta@unap.cl

**Lucía Ramos**

Catholic University of the North, Chile  
luramos@ucn.cl

**Igor Rižnar**

University of Primorska, Slovenia  
igor.riznar@guest.arnes.si

**Tania Tagle**

Catholic University of Temuco, Chile  
ttagle@uct.cl

**Lena Tica**

University of Kragujevac, Serbia  
lenaboki@gmail.com

**Vesna Tripković-Samardžić**

University "Mediterranean" Podgorica, Montenegro  
vesna.tripkovic-samardzic@fvu.me

**Alexey Tymbay**

Technical University of Liberec, Czechia  
alexey.tymbay@tul.cz

**Andrej Zavrl**

Slovenia  
andrej.zavrl@outlook.com

**Simon Zupan**

University of Maribor, Slovenia  
simon.zupan@um.si

# GUIDELINES FOR CONTRIBUTORS

## **ELOPE English Language Overseas Perspectives and Enquiries**

ELOPE publishes original research articles, studies and essays that address matters pertaining to the English language, literature, teaching and translation.

### **Submission of Manuscripts**

Manuscripts should be submitted for blind review in electronic form using the Faculty of Arts (University of Ljubljana) OJS platform (<https://journals.uni-lj.si/elope/about/submissions>). Only one contribution by the same author per volume will be considered. Each paper should be accompanied by abstracts in English and Slovene and keywords. Abstracts by non-native speakers of Slovene will be translated into Slovene by ELOPE. Please be sure to have a qualified native speaker proofread your English-language article. Suggested length of manuscripts is between 5,000 and 8,000 words.

### **Manuscript Style and Format**

The manuscript should be in the following format:

- title in English (no longer than 100 characters including spaces),
- abstracts in English and Slovene (with a maximum of 150 words for each language) and (up to five) keywords,
- the text should be divided into introduction; body of the paper (possibly subdivided); and conclusion.

The text should be preferably written in Word format (OpenOffice and RTF files are also acceptable). Please observe the following:

- 12-point Times New Roman font size,
- 2.5 cm page margins on all sides,
- 1.5-point line spacing,
- left text alignment,
- footnotes should be brief (up to 300 words per page; 10-point Times New Roman font size).

For resolving practical style and formatting queries, please see the articles in the latest on-line issue or contact the technical editor.

### **References**

References should comply with *The Chicago Manual of Style* (17th edition, 2017) author-date system. A Quick Guide to CMS is available here: [https://www.chicagomanualofstyle.org/tools\\_citationguide/citation-guide-2.html](https://www.chicagomanualofstyle.org/tools_citationguide/citation-guide-2.html).

### **Final note**

Please note that only manuscripts fully adhering to the ELOPE Guidelines for Contributors will be considered for publication.

# E L O P E

nglish  
anguage  
verseas  
erspectives and  
enquiries

**ELOPE Vol. 21, No. 1 (2024)**

## **Guest Editors**

Bernarda Leva, University of Maribor, Slovenia  
Simon Zupan, University of Maribor, Slovenia  
Nastja Prajnc Kacijan, University of Maribor, Slovenia

## **Journal Editors**

Smiljana Komar, University of Ljubljana, Slovenia  
Mojca Krevel, University of Ljubljana, Slovenia

## **Editorial Board**

Lisa Botshon, University of Maine at Augusta, United States of America; Biljana Čubrović, University of Belgrade, Serbia; Michael Devine, Acadia University, Canada; Dušan Gabrovšek, University of Ljubljana, Slovenia; Michelle Gadpaille, University of Maribor, Slovenia; Meta Grosman, University of Ljubljana, Slovenia; Allan James, University of Klagenfurt, Austria; Victor Kennedy, University of Maribor, Slovenia; Bernhard Kettemann, University of Graz, Austria; Alberto Lázaro, University of Alcalá de Henares, Spain; J. Lachlan Mackenzie, VU University Amsterdam, Netherlands; Tomaž Onič, University of Maribor, Slovenia; Roger D. Sell, Åbo Akademi University, Finland; Andrej Stopar, University of Ljubljana, Slovenia; Rick Van Noy, Radford University, United States of America; Terri-ann White, University of Western Australia, Australia

## **Editorial Secretary**

Gašper Ilc, University of Ljubljana, Slovenia

## **Technical Editor**

Andrej Stopar, University of Ljubljana, Slovenia

## **Proofreading**

Paul Steed

## **Editorial Policy**

*ELOPE: English Language Overseas Perspectives and Enquiries* is a double-blind, peer-reviewed academic journal that publishes original research articles, studies and essays that address matters pertaining to the English language, literature, teaching and translation.

The journal promotes the discussion of linguistic and literary issues from theoretical and applied perspectives regardless of school of thought or methodology. Covering a wide range of issues and concerns, ELOPE aims to investigate and highlight the themes explored by contemporary scholars in the diverse fields of English studies.

# **E**nglish **L**anguage **O**verseas **P**erspectives and **E**nquiries

## **Published by**

University of Ljubljana Press

*Založba Univerze v Ljubljani*

For the Publisher: Gregor Majdič, Rector of the University of Ljubljana

## **Issued by**

Slovene Association for the Study of English

*Slovensko društvo za angleške študije*

Department of English, Faculty of Arts, University of Ljubljana

*Oddelek za anglistiko in amerikanistiko, Filozofska fakulteta, Univerza v Ljubljani*

Ljubljana University Press, Faculty of Arts

*Znanstvena založba Filozofske fakultete Univerze v Ljubljani*

For the Issuer: Mojca Schlamberger Brezar,

Dean of the Faculty of Arts, University of Ljubljana

The journal is published with support from the Slovenian Research and Innovation Agency.

The publication is free of charge.

## **Universal Decimal Classification (UDC)**

Kristina Pegan Vičič

## **Journal Design**

Gašper Mrak

## **Cover**

Marjan Pogačnik: *Zimsko cvetje*, 1994

7.6 x 10.0 cm; colour etching, deep relief

Owner: National Gallery, Ljubljana,

Photo: Bojan Salaj, National Gallery, Ljubljana

## **Printed by**

Birografika Bori

## **Number of Copies**

110



**UNIVERSITY  
OF LJUBLJANA**

**FF**

**Faculty  
of Arts**

<https://doi.org/10.4312/elope.21.1>

Online ISSN: 2386-0316

Print ISSN: 1581-8918



This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License.