



Institute of Croatian Language and Linguistics
University of Rijeka • Faculty of Law

International Conference

*Terminology and Specialized Knowledge Representation:
New Perspectives on User Needs*

BOOK OF ABSTRACTS

University of Rijeka, Faculty of Law
Rijeka, 9–10 June 2022



International Conference

Terminology and Specialized Knowledge Representation: New Perspectives on User Needs

Book of Abstracts

Publisher

Institute of Croatian Language and Linguistics

Edited by

Ana Ostroški Anić

Larisa Grčić Simeunović

Ivanka Rajh

Corrections

Maja Matijević

Design

Davor Milašinčić

ISBN: 978-953-8390-10-4

PROGRAMME COMMITTEE

Martina Bajčić (Croatia), Łucja Biel (Poland), Ivana Brač (Croatia), Maja Bratanić (Croatia), Elena Chiocchetti (Italy), Rute Costa (Portugal), Thierry Declerck (Germany), Jorge Gracia (Spain), Larisa Grčić Simeunović (Croatia), Dagmar Gromann (Austria), Lana Hudeček (Croatia), Ilan Kernerman (Israel), Maja Lončar (Croatia), Patricia Martín-Chozas (Spain), Milica Mihaljević (Croatia), Emilia Mišćenić (Croatia), Liudmila Mockienė (Lithuania), Elena Montiel-Ponsoda (Spain), Bruno Nahod (Croatia), Henrik Nilsson (Sweden), Marija Omazić (Croatia), Petya Osenova (Bulgaria), Ana Ostroški Anić (Croatia), Fernando Prieto Ramos (Switzerland), Sigita Rackevičienė (Lithuania), Ivanka Rajh (Croatia), Marko Robnik-Šikonja (Slovenia), Susan Šarčević (United States), Kristina Štrkalj Despot (Croatia), Dimitar Trajanov (North Macedonia), Mojca Žagar Karer (Slovenia)

ORGANISING COMMITTEE

Ana Ostroški Anić (President), Martina Bajčić, Ivana Edmonds, Katja Dobrić Basaneže, Lobel Filipić, Maja Matijević, Martina Pavić

FINANCIAL SUPPORT

Croatian Science Foundation (HRZZ) under the installation research project HRZZ-UIP-2017-05-7169 Dynamicity of Specialized Knowledge Categories (DIKA)

European Association for Terminology (EAFT)

With the support from the COST action Nexus Linguarum: CA18209 – European network for Web-centred linguistic data science.

Contents

Plenary lectures

Špela Vintar: Frames, knowledge models and cognitive spaces: Do they overlap?

Friedemann Vogel: How to make the law more understandable – for whom? Results from an empirical evaluation of the linguistic legal editorial in the German Federal Ministry of Justice

Sara Carvalho: Terminology and Linguistic Linked Data in the Biomedical Domain: Hype, Hope and Reality

Presentations in sessions

Martina Bajčić, Emilia Miščenić: The Conundrum of Transparency in the EU Multilingual Digital Single Market

Sandra Cuadrado Camps, Bárbara L. Monllor Taltavull: The terms and the concepts behind the culture of consent

Thierry Declerck: Towards a Multilingual ATC Ontology

Kristina Despot, Liudmila Mockienė, Petya Osenova, Ana Ostroški Anić: Covid-19 and health related metaphors in Bulgarian, Croatian and Lithuanian ParlaMint corpora

Katja Dobrić Basaneže, Martina Bajčić: Terminology in the “New Normal”: A Corpus-based Study into Term Variation in Covid-19-related EU Legal Documents

Ivana Edmonds, Alenka Šuljić Petrc: English for Specific Purposes: Students’ Attitudes towards Professional Terminology and Teaching Approaches

Selma Elkasović, Jasmina Jelčić Čolakovac: Authentic vs. Instructive: The Representation of Specialised Maritime English Lexis in Authentic Teaching Materials

Christian Galinski, Blanca Stella Giraldo Perez: Minimal Interoperability Mechanisms (MIM) Requirements for Terminological and Lexicographical Data

Larisa Grčić Simeunović, Ivanka Rajh: Towards the Conceptual Organisation of the Interdisciplinary Domain of Smart Urban Coastal Sustainability (EU-CONEXUS)

Barbara Heinisch: Terminological Usability – Adapting Terminological Databases to Different User Groups According to Usability Principles

Zrinka Jelaska, Sara Košutar: Sound Acquisition Terms in Linguistics and Language of General Public

Lana Hudeček, Milica Mihaljević, Irena Miloš: Coronavirus-related terminology in the context of the European Union

Kristian Lewis, Marin Mrčela: Croatian legal terminology and Clear Writing movement: do citizens really understand judicial decisions?

Željka Macan, Luka Medak: Terminology in the Manual Official International Rules for Basketball

Marko Markovikj, Jovana Dobрева, Mary Lucas, Irena Vodenska, Lou Chitkushev, Dimitar Trajanov: Terminology and topics analysis of tweets related to the COVID-19 pandemic

Ana Meštrović, Slobodan Beliga, Mihaela Matešić: Automatic Detection of Terminology and Topics from COVID-19 Related Texts Published on Social Media

Ana Ostroški Anić, Ivana Brač: The function of semantic roles in linking general and specialised lexical resources

Martina Pavić: A corpus analysis of adjectives in different registers of Croatian medical terminology

Sigita Rackevičienė, Liudmila Mockienė: Terminology Studies at Mykolas Romeris University

Siniša Runjaić, Ana Šverko: From Initial Terminological Project to a Richly Illustrated Vocabulary: The Development of the Croatian Glossary of Classical Architecture

Introduction

The international conference *Terminology and Specialized Knowledge Representation: New Perspectives on User Needs* is organized within the research project HRZZ-UIP-2017-05-7160 *Dynamicity of Specialized Knowledge Categories* (DIKA), financed by the Croatian Science Foundation (HRZZ) from 2018 to 2023. The project focuses on using semantic frames to define specialized knowledge categories and identify links between general and specialized knowledge.

The conference aims to bring together specialists in the disciplines of terminology, terminography, specialized lexicography, LSP, computational linguistics and NLP in order to share their insights on the design and development of terminological materials, resources and applications with a particular view of emerging user needs in linking general and specialized knowledge resources.

The conference is organized around several main research topics:

1. changes in terminology work and specialized languages due to the pandemic and post-pandemic circumstances (handling neologisms; crisis terminology work)
2. terminology for non-specialists (creating new and/or adapting existing terminological resources intended for non-experts or the general public; adapting terminological definitions, identifying and describing synonyms, using different registers; specialized knowledge in educational materials and in teaching activities)
3. developing legal literacy (defining legal concepts in non-expert materials and resources, e.g. in user manuals, end user agreements, etc.)
4. terminology and specialized translation (including terminology work and machine translation; quality control)
5. terminology and linguistic linked open data in biomedical and life sciences domains
6. issues in linking general and specialized terminological resources (identification and presentation of semantic frames, multi-word terms and figurative units).

A workshop on terminology and linguistic linked open data (LLOD) is organized as a special conference session devoted to public health and biomedical sciences terminologies, with six papers included in the session. The idea for the workshop originated within the Task 4.4. *Life Sciences* of the COST action Nexus Linguarum, CA18209 – European network for Web-centred linguistic data science.

We would like to thank the Institute of Croatian Language and Linguistics and the Faculty of Law, University of Rijeka, for providing the support necessary for organising the conference. Thanks are also due to the European Association for Terminology (EAFT) for generously supporting the conference, as well as to Nexus Linguarum for being the backbone of cooperation in an international community of experts in the fields of terminology and linguistic linked data.

Finally, we thank all members of the Programme and Organising Committees for their hard work and efforts put into making this conference happen. We wish you successful and interesting presentations and discussions, and hope to see you all again.

Ana Ostroški Anić

President of the Organising Committee

PLENARY LECTURES

PLENARY LECTURE Thursday, 9 June 2022

Špela Vintar

Department of Translation Studies, Faculty of Arts, University of Ljubljana

spela.vintar@ff.uni-lj.si

Frames, knowledge models and cognitive spaces: Do they overlap?

Frames have gained popularity in terminology because of their potential of reconciling the cognitive, textual and formal views of specialised knowledge. As a result, the number of terminological knowledge bases employing the frame-based approach is growing, with emerging synergies between semantic web technologies and terminological knowledge resources.

With recent advances in Natural Language Processing, new methods of knowledge modelling and knowledge mining are being explored, hoping to facilitate the construction of linked specialised resources.

In the first part of the talk, we present the construction of a new knowledge base for karstology which employs the frame-based approach, combines manual and NLP methods to identify elements of knowledge, and includes graphs, images and maps into the representation of concepts.

In the second – more experimentally oriented – part, we concern ourselves with the question whether semantic relations as stipulated through frames, and semantic similarity as represented through nearest neighbours in word embeddings, coincide or overlap with cognitive spaces as attested through free word associations. We present the results of two experiments where the psycholinguistic method of free associations was used to collect responses from expert and lay participants. In the first experiment, we analyse the responses in search of implicit frames, and in the second, we compare them to the semantically similar words computed through word embeddings.

We conclude by interpreting the interesting findings from a multidisciplinary perspective, leaving plenty of room for discussion.

PLENARY LECTURE Thursday, 9 June 2022

Friedemann Vogel

University of Siegen

friedemann.vogel@uni-siegen.de

**How to make the law more understandable – for whom?
Results from an empirical evaluation of the linguistic legal editorial in the
German Federal Ministry of Justice**

It is an old question both in law and linguistics: how can we improve a text of legal norms and provide a better basis for the practice of legal interpretation not only in law but also in non-legal contexts? The main problem with this question and the answers given so far is that there are mostly only hypothetical assumptions about the addressees and actual recipients of legal texts. The lecture will present some results of an empirical project evaluating linguistic legal editing in German legislation (2019-2021), which will provide insight into the opportunities and challenges of developing a more comprehensible and addressee-oriented law.

PLENARY LECTURE Friday, 10 June 2022

Sara Carvalho

Águeda School of Technology and Management, University of Aveiro / NOVA FCSH
sara.carvalho@ua.pt

**Terminology and Linguistic Linked Data in the Biomedical Domain:
Hype, Hope and Reality**

Over the last few decades, healthcare has been impacted by a series of critical changes, mostly from a technological standpoint, which, in some cases, have led to hurdles concerning the dissemination of clinical data across multiple providers, institutions, countries, and stakeholders. With its interdisciplinary nature, Terminology can make a valuable contribution in this regard by supporting a more effective knowledge representation, organisation, and sharing in this field, thereby enhancing specialised communication among these various stakeholders. More specifically, it is believed that the potential resulting from the combined efforts of a double-dimensional (conceptual and linguistic) framework of terminology work and the more recent Linguistic Linked Data paradigm can help address some of the current challenges, namely those inherent to semantic interoperability, while also paving the way for new opportunities in linking different resources and making them available to both human and machine end users.

PRESENTATIONS IN SESSIONS

Martina Bajčić

Emilia Mišćenić

University of Rijeka, Faculty of Law

martina.bajcic@pravri.hr, emiscenic@pravri.hr

The Conundrum of Transparency in the EU Multilingual Digital Single Market

Undoubtedly, the channel of the law is changing in the digital age as legal information is available on websites, while disputes can be resolved by means of online forms. This digital turn is especially visible in the EU digital single market and business-to-consumer relationships (hereinafter: B2C relationships) in which information must be provided to consumers ‘in a clear and comprehensible manner’ in order to guarantee transparency (Mišćenić 2021 and Loos 2017). Only if consumers have been notified of the “essential” information of a given transaction, can they make an “informed choice” when entering a B2C relationship (e.g. Mišćenić 2018, 2021). As confirmed by the Court of Justice of the EU (hereinafter: CJEU), “information, before and at the time of concluding a contract, on the terms of the contract and the consequences of concluding it is of *fundamental importance* for a consumer” (Radlinger and Radlingerová, para. 64).

Indeed, the latter information duty contributes to the realisation of transparency which in turn highlights the need to frame information in plain, clear and comprehensible language. However, if the above essential information, which is necessary for an average consumer to make an informed choice, is not phrased in clear and consistent terms, the requirement of transparency is undermined (Mak 2020). Consequently, consumers enter B2C legal transactions ‘without understanding that they lack understanding’ of crucial information on possible outcomes and economic consequences of a transaction (Mišćenić 2018).

Despite the fact that the importance of transparency and information duty in the context of EU consumer law has been recognized (e.g. Mišćenić 2021), there have been no comprehensive studies into the role of language and terminology in achieving transparency. The present study attempts to fill this niche by analysing (multilingual) terms used to denote uniform concepts of EU consumer law observed against the principles of equal authenticity and continuity (Bajčić 2020, 2021). It will include both a corpus-informed and legal analysis in order to identify the link between the usage of inconsistent terminology in the relevant legislation and the requirement of transparency in EU law. In view of the fact that the “information model” crystalized in settled CJEU case-law, EU directives and the consumer *acquis* in general is difficult to translate into real life and B2C transactions, special emphasis is put on the information duty observed through the

lens of terminology used to denote key concepts of EU consumer law both in EU and national legislative texts.

Keywords: the requirement of transparency, information duty, EU digital single market, multilingualism, EU terminology.

References

- Bajčić, Martina. 2021. Making a corpus-linguistic U-turn in multilingual adjudication. *In Law, Language and the Courtroom*. Eds. Gozdz Roszkowski, Stanislaw; Pontrandolfo, Gianluca. Routledge. 85–97.
- Bajčić, Martina. 2020. Linguistic Comparison within CJEU's Decision-Making: A Debunking Exercise. *International Journal for the Semiotics of Law - Revue internationale de Sémiotique juridique*. 1–17.
- Loos, Marco B. M. 2017. Double Dutch: On the role of the transparency requirement with regard to the language in which standard contract terms for B2C-contracts must be drafted. *Journal of European Consumer and Market Law (EuCML)* 2. 54–59.
- Mak, Vanessa. 2020. *Legal Pluralism in European Contract Law*. Oxford Studies in European Law. Oxford University Press.
- Mišćenić, Emilia. 2021. Učinkovitost i transparentnost online informiranja potrošača. *Pravna zaštita pojedinaca na jedinstvenom digitalnom tržištu Europske unije*. Ed. Barbić, Jakša. Hrvatska akademija znanosti i umjetnosti. Zagreb. 39–63.
- Mišćenić, Emilia. 2018. Uniform Interpretation of Article 4(2) of UCT Directive in the Context of Consumer Credit Agreements: Is it possible? *Revue du droit de l'Union européenne* 1/3. 1–33.

Sandra Cuadrado Camps¹

Bárbara L. Monllor Taltavull²

¹ TERM CAT, ² Gender perspective legal expert and member in the association of Dones Juristes (women's jurists)

scuadrado@termcat.cat, monllortaltavull@gmail.com

The terms and the concepts behind the culture of consent

The term 'sexual harassment' was coined for the first time in 1974 by Mary Rowe to describe misconduct by men at Massachusetts Institute of Technology. Following this, a group of women addressed the same problem at Cornell in 1975. The term 'date rape' was first used by Susan Brownmiller in 1975. In her essay *The Mother of all Questions* (2017), Rebecca Solnit wrote that what is unsayable becomes unknowable and vice versa, until something breaks. New recognitions required new language, and feminism coined a plethora of terms to describe the individual experiences that the conversations of the 1960s and 1970s were beginning to flush out of hiding.

50 years later, violence against women is now a public issue. At last. *Sexual abuse, sexual assault, sexual harassment, stealthing...* Most of these terms already had Catalan denominations, but there were many conceptual problems, because they were often used in an ambiguous or contradictory way, and the legislative texts of different countries do not usually have definitions corresponding to the terminology they use. In addition, many concepts related to sexual offences have a significant cultural load, and they vary greatly depending on each country and each legislation which, in turn, usually changes periodically. For this reason, it was considered necessary to draw up a collection containing the basic terminology in this area with the principal aim of naming the violence against women, while focusing on the consent culture. To illustrate this flow approach, we will highlight the terminology co-creation process we have followed with our legal experts in order to develop the recent project on sexual consent.

The terminology includes more than 50 terms. The source of the terms on this project is a very wide range of literature: from activists' essays, media, legal references, conferences, relevant books about sexual violence against women and glossaries recommended by our team of legal experts, to social media lists and debates on feminism and the #metoo movement. Each term includes the denominations and the definition in Catalan, the equivalents in Spanish and English, and, in some cases, conceptual clarifications or complementary notes. This terminology dictionary compiles the terms related to sexual consent, key material for understanding that only yes is yes, and that any sexual practice must be voluntary, safe, desired, consensual, and free of violence.

Keywords: terminology, sexual consent, feminism, consent culture.

References

- Burgen, Stephen. 2018. Spain to introduce 'yes means yes' sexual consent law. *The Guardian*. www.theguardian.com/world/2018/jul/18/spain-to-introduce-yes-means-yes-sexual-consent-law (accessed 26 April 2022).
- Butler, Judith. 1993. *Bodies that matter: on the discursive limits of sex*. Routledge. London.
- European Institute for Gender Equality. 2022. *Glossary & Thesaurus*. Vilnius. <https://eige.europa.eu/thesaurus> (accessed 26 April 2022).
- GREVIO. 2022. Council of Europe Expert Group on Action against Violence against Women and Domestic Violence. Strasbourg. www.coe.int/en/web/istanbul-convention/grevio (accessed 26 April 2022).
- Istanbul Convention: Action against violence against women and domestic violence*. 2022. Council of Europe. Strasbourg. www.coe.int/en/web/istanbul-convention (accessed 26 April 2022).
- Khomami, Nadia. 2017. #MeToo: how a hashtag became a rallying cry against sexual harassment. *The Guardian*. www.theguardian.com/world/2017/oct/20/women-worldwide-use-hashtag-metoo-against-sexual-harassment (accessed 26 April 2022).
- Solnit, Rebecca. 2017. *The mother of all questions*. Haymarket Books. Chicago.
- The Consent Academy*. 2022. Seattle, WA. www.consent.academy (accessed 26 April 2022).

Thierry Declerck

DFKI, GmbH, Multilinguality and Language Technology, Saarland
Informatics Campus, Saarbrücken
declerck@dfki.de

Towards a Multilingual ATC Ontology

We describe the work consisting in building a multilingual ontology based on the Anatomical Therapeutic Chemical Classification System (ATC), with the goal of publishing it on the Linked Data cloud. Supervised by the World Health Organization Collaborating Center for Drug Statistics Methodology (WHOCC),¹ ATC offers a hierarchically organised set of codes that describe “active ingredients of drugs according to the organ or system on which they act and their therapeutic, pharmacological and chemical properties”.² This classification system aims at improving the quality of drug prescription and use.

There already exists a Linked Data compliant version of the ATC system, and it is available at <https://bioportal.bioontology.org/ontologies/ATC>. This (simple) ontology is also offering some links to related health information system (like, for example, via the property *umls:hasSty* to the UMLS Metathesaurus³). Figure 1 displays a screenshot of the ATC site within the BioPortal information offer.⁴

¹ See https://www.whooc.no/atc_ddd_methodology/purpose_of_the_atc_ddd_system.

² Quoted from https://en.wikipedia.org/wiki/Anatomical_Therapeutic_Chemical_Classification_System.

³ UMLS stands for “Unified Medical Language System” and is a metathesaurus for biomedical information. See https://www.nlm.nih.gov/research/umls/knowledge_sources/metathesaurus/index.html for more details.

⁴ The ontology can also be downloaded, as a *.ttl file, from this web page.

The screenshot shows the BioPortal interface for the ATC ontology. The left sidebar contains a hierarchical tree of classes, with 'ANABOLIC STEROIDS' selected. The right sidebar shows the 'Details' tab for this class, displaying the following information:

Preferred Name	ANABOLIC STEROIDS
ID	http://purl.bioontology.org/ontology/ATC/A14A
ATC LEVEL	3
cu	C0002845
Is Drug Class	Y
notation	A14A
prefLabel	ANABOLIC STEROIDS
tui	T125 T121
subClassOf	ANABOLIC AGENTS FOR SYSTEMIC USE

Figure 1: Screenshot of the ATC ontology in the BioPortal. Links to related resources are included in the “mappings” section of the page. Taken from <https://bioportal.bioontology.org/ontologies/ATC/?p=classes&conceptid=http%3A%2F%2Fpurl.bioontology.org%2Fontology%2FATC%2FA14A>

The BioPortal ATC ontology covers only the English language. It does not include definitions or notes to give some human readable indications about the topic of an ATC code.

We are extending the language coverage of this ontology, currently for German, Spanish, Hungarian and Polish, following different approaches for some of the considered languages. As we have access to the official German version of ATC, as an Excel file, through the German provider of statutory health insurance AOK,⁵ we transformed the CSV format of this file into RDF, and merged the resulting data with the English version, by just adding the German labels. In doing so, we automatically link the German labels to the UMLS Metathesaurus, when those links are also present in the Bioportal version of the ATC system. Figure 2 shows a (very small) part of the AOK Excel file.

⁵ The data can be downloaded here: <https://www.wido.de/publikationen-produkte/arszneimittel-klassifikation/amtliche-atc-klassifikation/>.

	A	B	C
1	ATC-Code	ATC-Bedeutung	
2			
3	A	ALIMENTÄRES SYSTEM UND STOFFWECHSEL	
4			
5			
6	A01	STOMATOLOGIKA	
7			
8			
9	A01A	STOMATOLOGIKA	
10			
11	A01AA	Mittel zur Kariesprophylaxe	
12	A01AA01	Natriumfluorid	
13	A01AA02	Natriummonofluorphosphat	
14	A01AA03	Olafur	
15	A01AA04	Zinn(II)-fluorid	
16	A01AA05	Fluorsilan	
17	A01AA30	Kombinationen	
18	A01AA51	Natriumfluorid, Kombinationen	

Figure 2: The Excel file with the German labels for the ATC codes

We should stress here that there is more information in the ATC file delivered by AOK, as for example the typical doses of a medicament to be administered. We do not consider this information for now.

We applied a similar transformation approach for the official Spanish data.⁶ The structure of the Excel file is different, but as we do need to extract only the ATC codes and the corresponding labels, it was straightforward to include the Spanish denominations as label values in our ontology. Figure 3 displays (an excerpt of) the Spanish Excel file. Here also, we stress the fact that the Excel file contains more information, like statistics on the administration of the drugs to different genders, age structures etc., to be considered in a next phase.

⁶ Available at the Spanish Ministry of Health: www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&ved=2ahUKEwjw9nVpZP2AhXCxqQKHdRA4EQFnoECDAQAQ&url=https%3A%2F%2Fwww.mscbs.gob.es%2FestadEstudios%2Festadisticas%2Festadisticas%2FestadMinisterio%2FSIAP%2FMedicamentos_BDCAP_2013.xls&usg=A0vVaw03z3OE7IGvA9UojNpmHcQX.

1	1	A	APARATO DIGESTIVO Y METABOLISMO
2	2	A01	PREPARADOS ESTOMATOLOGICOS
3	3	A01A	PREPARADOS ESTOMATOLOGICOS
4	4	A01AB	ANTIINFECCIOSOS DE USO ORAL LOCAL
5	5	A01AB09	MICONAZOL
6	5	A01AB11	VARIOS ANTIINFECCIOSOS ORAL LOCAL
7	5	A01AB19	PERBORATO SODICO
8	5	A01AB91	NISTATINA
9	4	A01AC	CORTICOSTEROIDES DE USO ORAL LOCAL
10	5	A01AC03	HIDROCORTISONA
11	5	A01AC53	HIDROCORTISONA EN ASOCIACION
12	4	A01AD	OTROS PREPARADOS DE USO ORAL LOCAL
13	5	A01AD02	BENCIDAMINA
14	5	A01AD11	VARIOS (TRATAMIENTO ORAL LOCAL)

Figure 3: The Excel file with the Spanish labels for the ATC codes

As for Hungarian and Polish, we first had a look at the corresponding Wikidata pages for ATC,⁷ which are already in a Linked Data compliant format and pointing to Wikipedia as the main source of information. Currently, we do not know if there are official pages displaying those codes in those languages (and for other languages listed in Wikidata, but with an uneven coverage of ATC categories). For now, we covered only the two levels of ATC by including those labels. Figure 4 displays a screenshot from our ontology editor, in which we can see how the multilingual coverage is achieved, where we link to the Wikidata and Wikipedia. And links to UMLS are automatically established for all those labels.

⁷ See <https://www.wikidata.org/wiki/Q192093>.

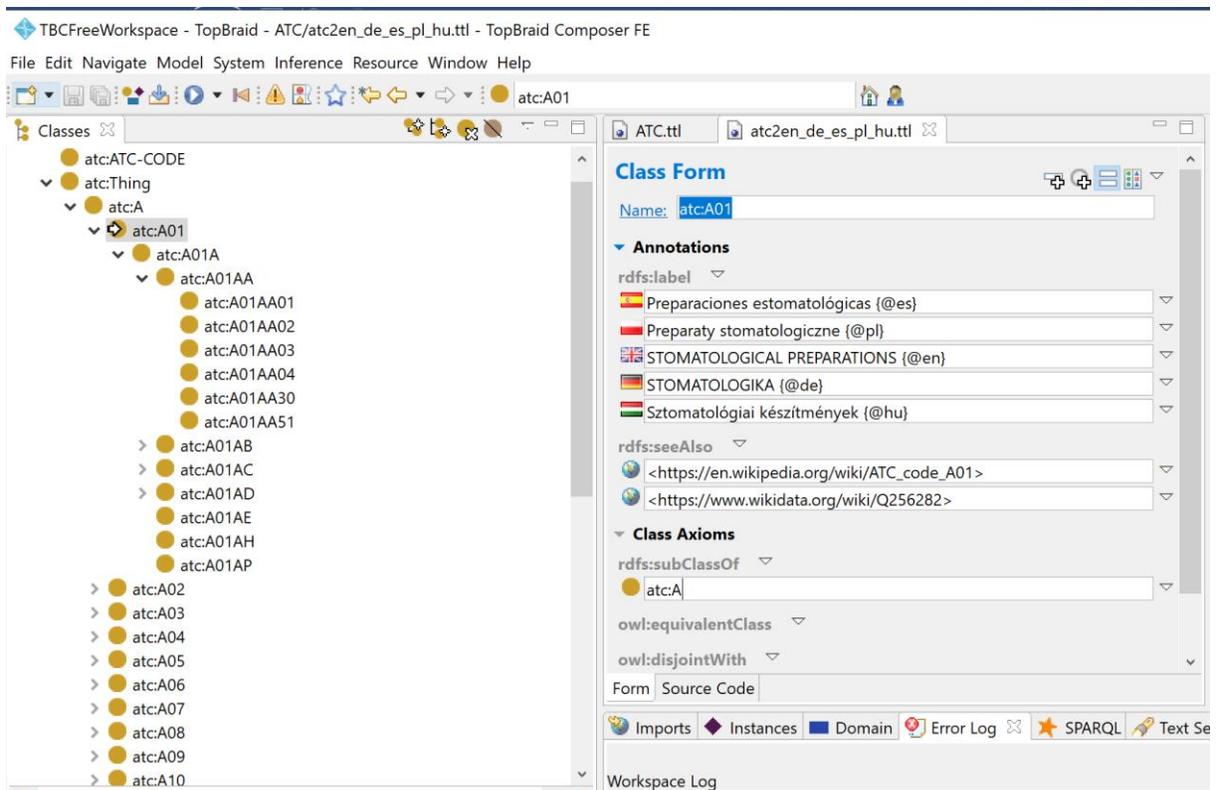


Figure 4: A screenshot from our multilingual ontology

Current Work

Current work is dealing with the connection of the values of the `rdfs:label` property to full lexical description of the terms, using for this OntoLex-Lemon (Cimiano et al. 2016), also for marking possible variants of each term. This step can possibly support improved terminology translation in the context of automated translation.

We are checking the multilingual coverage that was generated by our approach, as the source data can be from different versions of ATC. We are also starting to add (possible) definitions to the entities of our ontology. This can happen via the Wikidata Query Service, to get the Wikipedia pages associated with each listed ATC code. We can then, with the help of the Wikipedia API, extract from the Wikipedia articles the summary, which approximately could correspond to a definition, and which should be curated. In this way, we can offer a compact (in one file) multilingual ACT ontology that could be used as a multilingual terminology.

Acknowledgments

The work described in this contribution is supported by the Horizon 2020 research and innovation programme with the projects Prêt-à-LLoD (grant agreement no. 825182). It is based on discussion

within the COST Action NexusLinguarum – European network for Web-centered linguistic data science (CA18209), supported by COST (European Cooperation in Science and Technology).

We would also like to thank the anonymous reviewers for their helpful comments on the submitted version of this contribution.

References

Cimiano, Philipp; McCrae, John P.; Buitelaar, Paul. 2016. *Lexicon Model for Ontologies: Community Report*. W3C Community Group Final Report. www.w3.org/2016/05/ontolex.

Kristina Despot¹

Liudmila Mockienė²

Petya Osenova³

Ana Ostroški Anić¹

¹ Institute of Croatian Language and Linguistics, ² Institute of Humanities, Mykolas Romeris University, ³ Sofia University “St. Kliment Ohridski”, Institute of Information and Communication Technologies

kdespot@ihjj.hr, liudmila@mruni.eu, petya@bultreebank.org, aostrosk@ihjj.hr

Covid-19 and health related metaphors in Bulgarian, Croatian and Lithuanian ParlaMint corpora

The Covid-19 pandemic has been an all-pervading topic in all types of public discourse for the past two years, much of which consisted of defining the novel virus and referring to it in terms of various metaphorical framings. Several studies have shown that the discourse around the pandemic uses the war framing heavily, both in public discourse and in social media (Despot and Ostroški Anić 2021, Semino 2021, Wicke and Bolognesi 2020), despite the fact that the use of the war metaphor in health communication discourse has been heavily criticised and discouraged (Hauser and Schwarz 2020). Since the Lithuanian and Bulgarian data were processed first, some insights for the metaphorical frames in Bulgarian parliamentary data in comparison to the situation in Lithuania were reported in Osenova 2021.

Although many studies of the language used to communicate about Covid-19 explored discourse and spoken language, including messages sent by influential politicians, none has yet analysed the topic of Covid-19 in political discourse structured and regulated in the form of parliamentary debates. As transcriptions of spoken language, records of parliamentary debates consist of rich and valuable sociodemographic metadata. The ParlaMint project gathered transcriptions of the sessions of 17 European national parliaments (Erjavec et al. 2022), the corpora of which were used for the analysis of metaphorical conceptualization of the Covid-19 pandemic and related health topics in Bulgarian, Croatian and Lithuanian. The Bulgarian corpus includes all publicly available plenary sessions, i.e. from October 2014 to July 2020, downloaded from the Bulgarian National Assembly. The Croatian corpus contains speeches of the 9th term of the Croatian parliament, from November 2016 to May 2020, while the collection of Lithuanian parliamentary debates transcripts was compiled in 2020.

The analysis for each language followed the model represented in Despot and Ostroški Anić (2021), adapted for the analysis conducted in the NoSketch Engine. First, most common target words related to the Covid-19 were identified in all three languages. Their concordances were then analysed, and all sentences containing metaphorical expressions of target words extracted. The source frame invoked and the conceptual metaphor behind each metaphorical expression was defined for each sentence.

This paper discusses the results of the comparison of metaphors related to Covid-19 and related health topics extracted from the three mentioned parliamentary corpora: Bulgarian, Croatian and Lithuanian. Most common metaphors and frames for each language are described, with the focus on different variations of the shared metaphors in each language. Special focus is given to the possibilities of using this data in further research and the development of models used for linking shared conceptual knowledge in different linguistic and contextual settings, such as ontologies. Future goals include linking linguistic and conceptual metaphors following the MetaNet metaphor typology, and developing an ontology of semantic frames related to Covid-19 and public health.

Keywords: conceptual metaphor, corpora, Covid-19, parliamentary corpora.

References

- Despot, Kristina; Ostroški Anić, Ana. 2021. A War on War Metaphor: Metaphorical Framings in Croatian Discourse on Covid-19. *Rasprave: Časopis Instituta za hrvatski jezik i jezikoslovlje* 47/1. 173–208. doi.org/10.31724/rihjj.47.1.6.
- Erjavec, Tomaž; Ogrodniczuk, Maciej; Osenova, Petya et al. 2022. The ParlaMint corpora of parliamentary proceedings. *Language Resources and Evaluation*. doi.org/10.1007/s10579-021-09574-0.
- Hauser, David J.; Schwarz, Norbert. 2020. The war on prevention II: Battle metaphors undermine cancer treatment and prevention and do not increase vigilance. *Health Communication* 35/13. 1698–1704. doi.org/10.1080/10410236.2019.1663465.
- Osenova, Petya. 2021. Metaphors of Pandemia in the Bulgarian Parliamentary Data: a Corpus Survey. *Linguistic issues* 2/2. Eds. Todorova, Bilyana; Padareva-Ilieva, Gergana. 279–286. Blagoevgrad: University Publishing House “Neofit Rilski”. // Осенова, Петя. 2021. Метафорите на пандемията в българската парламентарна реч: корпусно изследване. *Лингвистични проблеми* 2/2. Тодорова, Биляна; Падарева-Илиева, Гергана (отг. редактори). 279–286. Благоевград: Университетско издателство “Неофит Рилски”]

- Semino, Elena. 2021. "Not Soldiers but Fire-fighters" – Metaphors and Covid-19. *Health Communication* 36/1. 50–58. doi.org/10.1080/10410236.2020.1844989.
- Wicke, Phillip; Bolognesi, Marianna M. 2020. Framing COVID-19: How we conceptualize and discuss the pandemic on Twitter. *PLoS ONE* 15/9: e0240010. doi.org/10.1371/journal.pone.0240010.

Katja Dobrić Basanež¹

Martina Bajčić²

¹ Juraj Dobrila University of Pula, Faculty of Arts, ² University of Rijeka, Faculty of Law
kdobri@unipu.hr, martina@pravri.hr

Terminology in the “New Normal”: A Corpus-based Study into Term Variation in Covid-19-related EU Legal Documents

It is now a well-known fact that COVID-19 was first detected in China's Wuhan and the world was made aware of this fact on 31 December 2019 when the World Health Organization reported on several cases of pneumonia in that city. Soon, the virus spread to other countries and the situation demanded action from the authorities. The European Commission, for instance, presented to the Council on 27 February 2020 its estimation on the possible impacts of the virus upon the industry in the EU. What followed were regulations, communications, Council decisions and reports related to the regulation of COVID-related issues at the EU level. Not only did these issues concern public health, but they also included many other areas, e.g. budget, competition, consumers, to mention but a few.

As a result, all these documents abound in new terms representing concepts brought about by the pandemic. Some of these terms represent neologisms, but many terms have also acquired new meanings (e.g. *pobrisati*) (Jozić et al. 2020). Needless to say, in our case terminology standardisation was conducted in English as the language of the draft texts, but these terms also had to be created in other official languages. Although the terms tend to be consistent in the language of the draft version, translated law, especially at the very beginning of term importation, is marked by term variation and instability (Bertaccini et al. 2010). This holds especially true for multi-word terms, sometimes also referred to as term-forming patterns (Biel 2014), since, the more extended the term becomes, the more difficult it is to approach its translation.

Studies focusing on terminological and phraseological variation in parallel corpora of EU law (Vigier and Sánchez Ramos 2017, Prieto Ramos and Guzmán 2018, Biel et al. 2018, Sosoni 2018, Dobrić Basanež 2018) all call for terminology standardisation in source texts in order to satisfy the principle of continuity of EU terms. This paper is thus an attempt to investigate term variation in Croatian translations of key EU legal documents related to COVID-19.

The corpus designed for this purpose is a parallel one and consists of English and Croatian legal documents related to the pandemic in various areas of EU law. It includes sixty-three documents in each subcorpus adopted in the time period from 2020 to 2021. These concern public

health, agriculture, budget, consumers, customs, digital single market, economic and monetary affairs, employment and social policy, enterprise, external relations, external trade, food safety, internal market, justice, freedom and security, maritime affairs, regional policy, research and innovation, taxation and transport. The English subcorpus consists of 267,893 words and the Croatian one of 241,086 words. AlignFactory is used to create a .tmx file containing parallel versions of texts which SketchEngine recognizes as a parallel corpus.

The methodology includes investigating parallel concordances of the pandemic-related keywords and determining their variants and stability. Results suggest that Croatian equivalents reveal a tendency towards foreignization strategies (Venuti 1995), which is especially visible in the translation of multi-word terms (e.g. *certificate holders' COVID-19 vaccination – cijepljenje protiv bolesti COVID-19 nositelja potvrda*). Furthermore, Croatian versions also point to the instability of terms due to their many variants (e.g. *booster doses – doze za docjepljivanje, dodatne doze, doze docjepljivanja; coronavirus disease 2019 – koronavirusna bolest 2019, bolest prouzročena koronavirusom iz 2019.; EU Digital COVID Certificate – EU digitalna COVID potvrda, digitalna potvrda EU-a o COVID-u*). Although in some cases one term in English can display two different concepts (e.g. *health – zdravlje, zdravstvo*) (Hudeček and Mihaljević 2012), we also encounter instances where the term stands for only one concept, but is nevertheless rendered through different variants with different meanings. The term *lockdown*, for instance, is a case in point, since according to several English monolingual dictionaries, it has only one meaning, but displays two variants in the Croatian corpus (e.g. *ograničenje kretanja, uvođenje mjera ograničavanja*).

Clearly, this may cause serious repercussions. Since term variation remains a recurrent problem in EU law (Biel and Koźbiał 2020, Bajčić 2022), this paper will especially focus on the meaning of term variants and their effects upon the principle of terminological consistency in EU legal translation.

Keywords: Covid-19, Croatian, English, EU legal translation, parallel corpus, term variation.

References

- Bajčić, Martina. forthcoming. Terminological Variation and Conceptual Divergence in EU Law. *Handbook of Legal Terminology*. Eds. Biel, Ł.; Kockaert, H. J. John Benjamins. Amsterdam – Philadelphia.
- Bertaccini, Franco; Massari, Monica; Castagnoli, Sara. 2010. Synonymy and variation in the domain of digital terrestrial television. Is Italian at risk? *Terminology in everyday life*. Eds. Thelen, Marcel; Steurs, Frida. John Benjamins. Amsterdam – Philadelphia. 11–19.

- Biel, Łucja. 2014. *Lost in the Eurofog: The Textual Fit of Translated Law*. Peter Lang. Frankfurt am Main.
- Biel, Łucja; Biernacka, Agnieszka; Jopek-Bosiacka, Anna. 2018. Collocations of Terms in EU Competition Law: A Corpus Analysis of EU English Collocations. *Language and Law: The Role of Language and Translation in EU Competition Law*. Eds. Marino, Silvia; Biel, Łucja; Bajčić, Martina; Sosoni, Vilelmini. Springer. Cham. 249–274.
- Biel, Lucja; Koźbiał, Dariusz. 2020. How do translators handle (near-) synonymous legal terms? A mixed-genre parallel corpus study into the variation of EU English-Polish competition law terminology. *Estudios de Traducción* 10. 69–90.
- Dobrić Basanež, Katja. 2018. Binomials in EU Competition Law. *Language and Law: The Role of Language and Translation in EU Competition Law*. Eds. Marino, Silvia; Biel, Łucja; Bajčić, Martina; Sosoni, Vilelmini. Cham: Springer. 225–248.
- Hudeček, Lana; Mihaljević, Milica. 2012. Zdravlje i zdravstvo. *Jezik* 59. 15–20.
- Jozić, Željko; Hudeček, Lana; Mihaljević, Milica. 2020. Mrežni izvori u vrijeme korone. Pojmovnik koronavirusa. *Hrvatski jezik* 2. 20–24.
- Prieto Ramos, Fernando; Guzmán, Diego. 2018. Legal Terminology Consistency and Adequacy as Quality Indicators in Institutional Translation: A Mixed-Method Comparative Study, *Institutional Translation for International Governance: Enhancing Quality in Multilingual Legal Communication*. Ed. Prieto Ramos, Fernando. Bloomsbury. London. 81–101.
- Sosoni, Vilelmini. 2018. Language and Translation in EU Competition Law: Insights from English, Greek, Italian and Spanish Versions of Legislative Texts. *Language and Law. The Role of Language and Translation in EU Competition Law*. Eds. Marino, Silvia; Biel, Łucja; Bajčić, Martina; Sosoni, Vilelmini. Springer. Cham. 179–205.
- Venuti, Lorenzo. 1995. *The translator's invisibility. A history of translation*. Routledge.
- Vigier, Francisco J.; Sánchez Ramos, María del Mar. 2017. Using parallel corpora to study the translation of legal system-bound terms: The case of names of English and Spanish courts. *Europhras: Proceedings of Computational and corpus based phraseology*.

Ivana Edmonds¹

Alenka Šuljić Petrc²

¹ University of Rijeka, Faculty of Law, ² University of Rijeka, Faculty of Tourism and Hospitality Management, Department of Foreign Languages in Business
iedmonds@pravri.hr, alenkasp@fthm.hr

English for Specific Purposes: Students' Attitudes towards Professional Terminology and Teaching Approaches

Students' attitudes towards learning English for specific purposes are pivotal in increasing their motivation and unorthodox foreign language skills that contribute to the overall greater success in the English language knowledge. Once they graduate from the university, students need to be prepared for a competitive job market, which nowadays requires even greater mastery of English language skills, since Croatia is a part of the European Union.

The aim of this study is to identify - by means of a survey - the attitudes of twenty pedagogy freshmen students from the University of Rijeka toward learning English for specific purposes (ESP) and teaching methods used to obtain the wanted learning outcomes. The focus of the study is on the implications that arrive from the survey results in terms of important factors in English language learning, unorthodox language learning methods, professional terminology related to the field of pedagogy, and reading strategies practised on various professional articles and authentic texts, with the purpose to improve students' comprehension.

In this research, the existing theories on ESP teaching methods were analysed, pointing out the appropriate teaching methods for this course (English in Education): the learner-centred approach, and the communicative method. The learner-centred approach focuses on students' needs by creating content, teaching materials, and assessment (Hutchinson and Waters 1987). Furthermore, it implies that teachers should allow students to choose the topic of interest so they could become more involved (Richards 2001). The (authentic) materials selected were various pedagogy articles and excerpts from the books based on students' interests (their major), and their English proficiency levels. In addition, they involved videos and stories from real-life multicultural settings. The use of authentic materials proved to be more stimulating for students and for communicating more meaningful messages (Baghban and Ambigaparhy 2011). The communicative method was based on the needs analysis while focusing on real-life examples through intercultural communicative competence (Byram 2014). The teaching methods served as tools to prepare meaningful lessons on terminology, which were introduced in a form of specific expressions, vocabulary, and idioms that were used in the field of pedagogy based on task-based

learning (Ellis 2009). Students had to create their own dictionaries from the analysed materials using words/expressions of their choice and their own examples. Furthermore, they had to analyse an article of choice, using pre-taught reading strategies to assist them with the terminology acquisition process.

The survey was conducted to examine students' attitudes towards professional terminology they were introduced to in the area of pedagogy and towards teaching approaches that were used. Students were given a survey at the beginning and at the end of the semester. The survey was created as a combination of structured (scaled questions) and unstructured questions (open-ended questions that can express students' view on certain issues/feedback). Students' feedback contained short-answer responses providing their experiences and opinions on the given questions. For example, their opinion on the effectiveness of terminology learning, additional knowledge that was obtained, important factors when learning English, or application of the terminology to reading comprehension. The scaled questions used a 5-point Likert-scale (1 being 'completely disagree' and 5 being 'completely agree') (Tullis and Albert 2013). From their answers conclusions could be made regarding the learning experiences while learning terminology in pedagogy and the future preferences regarding similar courses in their study area. The survey results will contribute to the future teaching of ESP in the field of pedagogy and help determine the focus for future classes that would better fit students' needs and motivate them to improve essential language skills for their future professions.

It is expected that the findings will suggest students' positive attitudes towards learning English for specific purposes, since there is a need for the use of professional literature that is available in English. In addition, development of positive attitudes towards the field-of-study related terminology will help students in becoming more proficient in the academic language. Higher English proficiency is achieved through continuous practice and exposure to classes that are tailored to support the content learning in the English language. Such classes would offer the necessary support in both the content area and English language, achieving the ultimate goal of academic English proficiency.

The contribution of this research lies in possibilities for further research, which is envisioned to examine the attitudes of law students and hospitality and tourism management students towards English for specific purposes (ESP) that is offered at their faculties. Comparisons can be drawn between survey results for Legal English and English in Tourism courses, the reference courses of English for Specific Purposes (ESP). The results will serve as a learning tool for the better fitted English language courses.

Keywords: English for Specific Purposes (ESP), students' attitudes towards ESP, practical knowledge, professional terminology, English in Education.

References

- Baghban, Zahra Z. V.; Ambigapathy, Pandian. 2011. A Review on the Effectiveness of Using Authentic Materials in ESP Courses. *English for Specific Purposes World* 10. 1–14.
- Byram, Michael. 2014. *Teaching and assessing intercultural communicative competence*. Shanghai Foreign Language. ShangHai.
- Ellis, Rod. 2009. Task-based language teaching: sorting out the misunderstandings. *International Journal of Applied Linguistics* 19. 221–246.
- Hutchinson, Tom; Waters, Alan. 1987. *English for Specific Purposes A Learner-Centered Approach*. Cambridge University Press. New York.
- Richards, Jack C. 2001. *Curriculum Development in Language Teaching*. Cambridge University Press. Cambridge. dx.doi.org/10.1017/CBO9780511667220.
- Tullis, Tom; Albert, Bill. 2013. *Measuring the User Experience: Collecting, Analyzing, and Presenting Usability Metrics*. Elsevier Inc. Amsterdam. doi.org/10.1016/C2011-0-00016-9.

Selma Elkasović¹

Jasmina Jelčić Čolakovac²

¹ Juraj Dobrila University of Pula, ² Faculty of Maritime Studies, University of Rijeka
selmael95@gmail.com, jasmina.jelcic@pfri.uniri.hr

Authentic vs. Instructive: The Representation of Specialised Maritime English Lexis in Authentic Teaching Materials

It is believed that authentic texts in ESP (*English for Specific Purposes*) are more likely to engage and intrinsically motivate learners towards learning a foreign language. According to Tominac Coslovich and Borucinsky (2010), the educational value of authentic materials is in the fact that they give us insight into language change, as well as changes outside the classroom. Furthermore, Kelly et al. (2002) state that, through their usage of authentic materials for over 10 years, they have found that such materials result in a positive learning attitude. On the other hand, simplified, adapted materials present students with a language that is more manageable, which leads to successful learning (Velazquez 2007). Additionally, adapted materials provide a more accessible way of learning for students at all levels, since lower-level students might struggle with the understanding of authentic texts (Tominac Coslovich and Borucinsky 2010).

Choosing between authentic and adapted materials in ESP classrooms is especially important in vocabulary teaching. Many studies have shown positive effects of the usage of authentic materials on vocabulary retention, due to an increased students' interest in the target language (Berwald 1987). On the other hand, vocabulary presented in authentic materials can often seem unnecessary or redundant, due to the difficulty and complexity of the language in question (Richards 2001; cited in Kilickaya 2004).

The focus of the study was to investigate to what extent authentic materials can actually be used in the instruction of ESP lexis, namely Maritime English vocabulary. In the first research stage, authentic materials (AuMs) were contrasted against adapted texts (AdMs) on the same topic. A detailed text analysis was carried on the lexical level in order to gain insight on how specialised knowledge on the given topic is represented within AuMs and AdMs respectively. In the second stage, a carefully planned lesson was carried out with university-level learners of Maritime English in order to provide answers to the following research questions: *Are authentic and adapted materials equally efficient in the representation of specialised knowledge when they are used in vocabulary instruction? Are learners equally likely to retain new lexis when it is presented in authentic texts as opposed to adapted texts?*

The lesson carried out with university-level learners of Maritime English consisted of three stages: the pre-activity stage, the main-activity stage and the post-activity stage. In the pre-activity, a pre-test was conducted in order to determine that the vocabulary items tested were not already familiar to the participants. The pre-test consisted of 20 vocabulary items, for which the participants were asked to choose the appropriate definition in a multiple-choice question. Next, the participants were divided into two groups where Group 1 was presented with an authentic text in which the 20 vocabulary items were highlighted. Similarly, Group 2 was presented with an adapted text which contained the same highlighted vocabulary. For the main activity, the participants were instructed to read the texts again, focusing on the highlighted vocabulary and the context in which the vocabulary items were presented. They were encouraged to infer the meanings from the context or look the definitions up in an online dictionary. The participants' retention of new lexis was then tested by means of a cloze test that consisted of the same vocabulary items in the sentential context.

Our results showed that greater vocabulary retention was recorded with AdMs than with AuMs; the adapted text presented the participants with a simplified context, hence the tested vocabulary items were more easily retained and used in a new context.

Possible further research could include different types of tasks for testing vocabulary retention and usage. Furthermore, classroom organisation could be taken into consideration by studying the effect of group work on vocabulary learning.

Keywords: authentic materials, adapted materials, vocabulary instruction, specialised knowledge.

References

- Berwald, Jean Pierre. 1987. Teaching Foreign Languages with Realia and Other Authentic Materials. *ERIC Q&A*. 3–5. <https://files.eric.ed.gov/fulltext/ED289367.pdf> (accessed 25 February 2022).
- Kelly, Charles et al. 2002. Effective Ways to Use Authentic Materials with ESL/EFL Students. *The Internet TESL Journal* VIII/11. <http://iteslj.org/Techniques/Kelly-Authentic.html> (accessed 25 February 2022).
- Richard, Jack C. 2001. *Curriculum development in language teaching*. Cambridge. Cambridge University Press.
- Kilickaya, Ferit. 2004. Authentic Materials and Cultural Content in EFL Classrooms. *The Internet TESL Journal* X/7. <http://iteslj.org/Techniques/Kilickaya-AuthenticMaterial.html> (accessed 25 February 2022).

- Tominac Coslovich, Sandra; Borucinsky, Mirjana. 2010. Using authentic materials in teaching Maritime English. *Proceedings of the 22nd International Maritime English Conference*. 307–317. www.pfri.uniri.hr/bopri/IMEC_Proceedings/PDF/IMEC22.pdf (accessed 25 February 2022).
- Velazquez, Ashley C. 2007. The Use of Authentic Texts in the K-12 Spanish Program. *Studies in Teaching 2007 Research Digest; Research Projects Presented at Annual Research Forum*. 133–138. <https://files.eric.ed.gov/fulltext/ED503269.pdf> (accessed 13 April 2022).

Christian Galinski

Blanca Stella Giraldo Perez

International Information Centre for Terminology (Infoterm)

christian.galinski@chello.at, blancaese@gmail.com

Minimal Interoperability Mechanisms (MIM) Requirements for Terminological and Lexicographical Data

Content interoperability and related terms are standing for different, though overlapping concepts e.g., in connection with repositories of structured content holding different kinds of content entities. Structured content can be differentiated into a variety of apparently different but related microcontent entities (microCE). The latter are defined as digital objects at the level of lexical semantics with a high degree of granularity each standing for “one idea or concept”. In this sense, terminological and lexicographical entries are microCEs, which besides other kinds of microCEs emerged out of different scientific approaches or diverse applications.

Due to the development of information and communication technologies (ICT), there are more applications and increasingly new ones where microCEs must be reusable or even re-purposable in all kinds of human-human and human-machine communication, such as in user-interface design of technical devices/systems, data repositories in the public domain, eLearning (especially microlearning), etc. Already since around 1990, the industry has become aware of the need for multilingual and multimodal content (beyond written and spoken) and developed standards for localization (L10N). Given this situation, standards related to microcontent are necessary under two interrelated aspects: standardised microCEs and their standardised principles and methods to be standardised preferably at the international level. Thus, content unification, standardisation and harmonisation are increasing exponentially in all kinds of regulations.

The fields of terminology and librarianship were pioneering as from the 1970s with pertinent international standards. Today, any application that has or could have a user interface needs such standards. No wonder that hundreds of technical associations and technical committees exist which are developing content-related standards – numbers growing. Big application fields comprise all kinds of “smart” applications and eApplications in technical documentation and translation, product master data management, medical informatics, eAccessibility and eInclusion – dealing with linguistic and non-linguistic data. In many applications terminological and lexicographical data must be displayed side-by-side or even (seemingly or really) be combined.

Increasingly, microCEs fall under the category of “critical communication”, which on all accounts requires stricter standards and new standards under the requirements of “comprehensive

content interoperability” with high demands for content quality and enhanced content reusability. Given the growing needs for and, therefore, an exponential increase of all sorts of microCEs, approaches for ICT-supported cooperative content creation are mushrooming, triggering a demand for pertinent standards. The MIM (minimal interoperability mechanisms) approach provides clues for revising existing or developing new standards in the fields of terminology and lexicography that comply with current and future requirements – fully taking into account “comprehensive interoperability” as well as economical aspects. Terminological and lexicographical data and other kinds of microCEs can be made highly interoperable where and to the degree necessary if their different communicative roles are duly respected.

Keywords: minimal interoperability mechanisms (MIM), terminological and lexicographical data, microcontent, comprehensive content interoperability, eAccessibility and eInclusion, user interface design.

Larisa Grčić Simeunović¹

Ivanka Rajh²

¹ University of Zadar, Department of French and Francophone Studies, ² Department of Romance Languages, Faculty of Humanities and Social Sciences, University of Zagreb

lgrcic@unizd.hr, irajh@m.ffzg.hr

Towards the Conceptual Organisation of the Interdisciplinary Domain of *Smart Urban Coastal Sustainability (EU-CONEXUS)*

In this paper, we deal with the terminology of a new interdisciplinary domain of ‘smart urban coastal sustainability’ underlying the transnational European university project EU-CONEXUS⁸. The goal of the project is to offer a minor university program or a specialisation using the cross-disciplinary approach to different social, economic and cultural elements making up the coastal environment. Such alliance of experts and specialised knowledge requires the elaboration of the multilingual terminological database that will allow the researchers, and students in this case, to learn how to think between the disciplines and integrate them (Kluger and Bartzke 2020; Newell 2001; Repko 2007; Tress et al. 2005).

The challenge resides in the fact that, in the process of emergence of the new cross-disciplinary field, sources from the existing domains like ecology, economy, tourism, biology and geography are being used with the aim of creating an interdisciplinary conceptual system. A terminologist should bear in mind that disciplines included in the new domain share a set of common concepts, which they approach from different perspectives. As the borders of these domains are not clear-cut, it is inevitable that terms transgress from their source domains to neighbouring domains (Cabr e 1995; Schnieder and Wegener 2010).

Our analysis will be performed on a comparable English-French-Croatian corpus where for each language the corpus consists of two sub corpora, one containing scientific texts and the other didactic texts. A pilot study of the corpus showed the difficulties of determining the boundaries of the new interdisciplinary field and raised the question of the possibility of creating a new conceptual system. In this paper, we aim to investigate the challenges of a conceptual and terminological analysis for terms that have already been defined as part of neighbouring domains but are now presented inside a new interdisciplinary perspective. Special attention will be given to terms that change their meaning due to migration to a new domain.

⁸ www.eu-conexus.eu/en.

Our preliminary results show that the Croatian authors mostly take the economic perspective on coastal development while the French rather take the ecological one. Understanding the differences helps to understand the intertwining of different fields and the efforts of scientists to explain, for instance, the problem of climate change that simultaneously affects the economy, urban development, reorganisation of coastal areas and tourism. Previous research has also shown that not only the terminological definitions are important for the researchers to be able to communicate with each other, but also affect how concepts are formed and understood in a society. In our research, we will use a bottom-up approach to arrive at a list of conceptual categories by starting from available data in texts and dictionary entries. The top concepts will be arrived at through the analysis of definitions as well as through calculations of frequency in our corpus of specialised texts.

Keywords: conceptual organisation, interdisciplinary domain, comparable English-French-Croatian corpus.

References

- Cabré, M. Teresa. 1995. On diversity and terminology. *Terminology* 2/1. 1–16
- Kluger, Max Oke; Bartzke, Gerhard. 2020. A practical guideline on how to tackle interdisciplinarity. A synthesis from a post-graduate group project. *Humanities and Social Sciences Communication* 7:47. doi.org/10.1057/s41599-020-00540-9.
- Newell, William H. 2001. A theory of interdisciplinary studies. *Issues in integrative studies* 19. 1–25.
- Repko, Allen F. 2007. Integrating interdisciplinarity: How the Theories of Common Ground and Cognitive Interdisciplinarity Are Informing the Debate on Interdisciplinary Integration. *Issues in Integrative Studies* 25. 1–31.
- Schnieder, Lars; Wegener, Marco. 2010. *Transgression of semantic boundaries by methodical terminology management*. International Symposium on Spatial Accuracy Assessment vol. 9.
- Tress, Gunther; Tress, Bärbel; Fry, Gary. 2005. Clarifying Integrative Research Concepts in Landscape Ecology. *Landscape Ecology* 20. 479–493.

Barbara Heinisch

Centre for Translation Studies, University of Vienna, Austria

barbara.heinisch@univie.ac.at

Terminological Usability – Adapting Terminological Databases to Different User Groups According to Usability Principles

Terminology, by definition, belongs to a specialised language and is, therefore, intended for domain experts. However, non-specialists are occasionally also confronted with terminology from a wide range of specialised languages. Terminological databases are a means to access and manage this specialist knowledge. However, terminological databases are specialist tools as well. Therefore, non-experts may face difficulties when using terminological databases, not only content-wise but also as an application.

This paper addresses the concept and practical implementation of terminological usability. According to the related ISO standard (International Organisation for Standardisation ISO 9241-11:2018), usability is referring to the extent to which a product can be used by certain users to achieve goals effectively, efficiently and satisfactorily in a certain context. The usability of terminological databases is still an underexplored area in academic literature (Marcos et al. 2006; Sevriens 2010; Heinisch 2019) although usability plays a dual role in terminology. On the one hand, it can refer to the usability of a terminological database, i.e. the application, e.g. its functions and navigation. On the other hand, usability can also refer to the usefulness of the content of a terminological resource for a given user.

Therefore, this paper analyses the adaptation of an existing bilingual terminological database in an organisation to the needs of non-experts. In this case study, the managers of the terminological database are terminologists but none of the users are neither familiar with the basics of terminology work nor with the terminology of terminology. According to a needs assessment, the users just require ‘translations’, ‘definitions’ or ‘background information’. To ease the access to terminology for these users, usability principles are taken into account. Usability thus requires adaptations of the user interface of the terminological database, including different user interfaces for different user groups and different terminology, e.g. of functions and navigation in the user interface. Since the terminological database covers different domains in which the organisation is active, including human resources, finance as well as education, the usefulness of the terminological database’s content, i.e. the terminologies, for different user groups also plays a significant role. Here, adaptations may require a deviation from common terminological

principles, such as deviations from the good practice of writing definitions, the number of definitions provided or the visibility of sources.

During the first step all the terms actually used in the organisation were collected as part of descriptive terminology work. This descriptive terminology work served as the basis for subsequent prescriptive terminology work. The origin of all the terms was collected together with the terms in a 'source' field in the terminological database. Here, an ethical issue arose because the terminological database now contains terms that were previously coined by organisational units but are currently deprecated. In this case study, it was deliberately chosen to hide the provenance of the terms in order to not embarrass those organisational units that have used the deprecated term before.

Other adaptations related to terminological usability are the engagement of the users in the further development of the terminological database. While a small number of users also participated in the initial design of the terminology application, the majority of the users were introduced to the final application. Nevertheless, through feedback options users can directly contribute to the improvement of the terminological database content. These may include simple adjustments such as the correction of spelling mistakes or work-intensive contributions to systematic terminology work in a certain domain of the organisation, either monolingually or bilingually.

Further adaptations to increase the usability of the terminological database include different terminology workflows for different domain experts. While some domain experts approve and release terminology via predefined workflows in the terminological database itself, other domain experts prefer to discuss terminological issues with the terminologists in person. Although this increases the amount of follow-up work for the terminologists, it is nevertheless an important step to meet the different needs of the user groups.

In conclusion, creators of terminological resources can influence the terminological usability. This paper shows the challenges and possible solutions in making adaptations to terminological databases in order to strike a balance between usability requirements on the one hand, and the good practice of terminology management and interoperability of terminological data, on the other.

Keywords: usability, terminology workflows, non-terminologists, ethics.

References

Heinisch, Barbara. 2019. Darstellung und Verbreitung von Terminologie mittels Terminologiedatenbanken. *Terminologie: Epochen – Schwerpunkte – Umsetzungen: Zum*

25-jährigen Bestehen des Rats für Deutschsprachige Terminologie. Eds. Drewer Petra; Pulitano, Donatella. Springer. Berlin. 117–142.

International Organization for Standardization. 2018. *ISO 9241-11. Ergonomics of human-system interaction — Part 11: Usability: Definitions and concepts*. International Organization for Standardization. Geneva.

Marcos, Mari-Carmen et al. 2006. Usability evaluation of online terminology databases. *Hipertext.net* 4.

Sevriens, A.P.M. 2010. *Improving usability for a terminology search website*. Master thesis, Tilburg University, Tilburg.

Zrinka Jelaska¹

Sara Košutar²

¹ Department of Croatian Language and Literature, Faculty of Humanities and Social Sciences, University of Zagreb, ² Department of Speech and Language Pathology, University of Zagreb
zjelaska@ffzg.hr, sara.kosutar@erf.unizg.hr

Sound Acquisition Terms in Linguistics and Language of General Public

Language acquisition is an important theoretical and applied field of research (especially in general and clinical phonology, and psycholinguistics), but it is also a very popular topic for the general public, starting with the parents and relatives of infants and young children. Phonological development plays a very important role in the first years of a child's life, as this is a sensitive period that has been widely studied in phonology (e.g., Flege et al. 1999). In this paper, we discuss the terminology of sound acquisition in children, which consists of normal systematic changes that can affect an entire category of phonemes and then gradually disappear as children grow older. The aim is to identify and describe the relationships between terms that describe different phonological errors. Naming the phenomenon itself is challenging: *phonological processes* (e.g., Stampe 1979), *phonological patterns* (e.g., Hodson 2004; Bernthal et al. 2017), or *phonological errors* (e.g., Dodd et al. 2003), the latter is sometimes distinguished from the former and attributed to language delay or a phonological disorder.

The terms discussed in this paper include synonyms such as *reduplication* vs. *doubling*, *elision* vs. *deletion*, which can also consist of one word vs. two words or three words: *liquid simplification* vs. *deliquiding*, *consonant cluster reduction* vs. *cluster reduction*; polysemic names such as *hyperonym assimilation* and *assimilation* as one of the hyponyms; paronyms such as *duplication* vs. *reduplication* or *deletion* vs. *omission*; names that appear semantically transparent, e.g. *duplication*, *reduplication*, *stopping*; and those that are difficult to pronounce and memorise, e.g., *coalescence*. Thus, *duplication* can be a synonym for *reduplication* (and then it is redundant), it can be an academic synonym (a word more commonly used by scholars), but it can also be a paronym: *reduplication* is the repetition of a word or part of a word to form a new one ('to double, to do something again'), while *duplication* can exclude the first identical element.

Synonymous terms will also be discussed because they describe the same process from a different point of view or assign a different theoretical approach. For example, *stopping* means replacing fricatives and affricates with stops; *deaffrication* means replacing affricates with fricatives or stops; *alveolarization* means replacing non-alveolar sounds with alveolar sounds. If

the child pronounces *tips* instead of *chips*, or *too* instead of *shoe*, the question is whether this is *stopping*, *deaffrication*, or *alveolarization* and whether it is meaningful to non-specialists.

The differences between the scientific terms that have theoretical implications and the terms for the general public will also be discussed, both within a language and between languages. Sometimes it is problematic to translate national terminology into English as an international language, since not all terms developed in one language have a true English equivalent, e.g. *fonološka odstupanja*, lit. ‘phonological deviations’, should perhaps be better translated as ‘phonological inaccuracies’. And when new national terms are introduced, they need to be translated into English, as it is an international language of communication, e.g. *deliquidizing*, *deliquidizing*.

Keywords: sound acquisition, terminology, English, Croatian.

References

- Bernthal, John; Bankson, Nicholas W.; Flipsen, Peter, Jr. 2017. *Articulation and phonological disorders: Speech sound disorders in children*. Pearson. New York.
- Dodd, Barbara; Holm, Alison; Hua, Zhu; Crosbie, Sharon. 2003. Phonological development: A normative study of British English-speaking children. *Clinical Linguistics & Phonetics* 17/8. 617–643. doi.org/10.1080/0269920031000111348.
- Flege, James Emil; Yeni-Komshian, Grace H.; Liu, Serena. 1999. Age Constraints on Second-Language Acquisition. *Journal of Memory and Language* 41. 78–104.
- Stampe, David. 1979. *A Dissertation on Natural Phonology*. Doctoral dissertation. Department of Linguistics. University of Chicago.
- Hodson W., Barbara. 2004. *HAPP-3: Hodson Assessment of Phonological Patterns*. Pro-Ed. Austin, Texas.

Lana Hudeček

Milica Mihaljević

Irena Miloš

Institute of Croatian Language and Linguistics

lhudecek@ihjj.hr, mmihalj@ihjj.hr, idrpic@ihjj.hr

Coronavirus-related terminology in the context of the European Union

European legislation-related terminology brings new concepts that emerge, inter alia, due to a particular policy. These terms need to be processed and incorporated into the Croatian language system in the same way as they are integrated into other official languages of the European Union. It is a common European terminology whose concepts are created in European legislative institutions at the language-independent level (LIL) and are designed by policymakers, i.e., authors of regulations and clients of their translations into all official languages of the European Union. For the definitions of these concepts, the source language is usually English (for most institutions) or French for the Court of Justice of the European Union. The process of transposing concepts into EU languages involves central terminological coordination bodies of the institutions, which work closely with contracting authorities, on the one hand, and with terminologists in language and translation services, primarily in translation and interpretation departments, on the other hand. Terminology is thus an integral part of the translation process. It takes place in a multilingual and multicultural community with the support of national experts and academia in the Member States. The most important concepts and terms are entered into IATE (European Union Terminology / InterActive Terminology for Europe). It is a database with multimillion entries in twenty-five languages (EU languages and Latin), with structured metadata only in English and a term reliability label. Terms marked with one or two asterisks have low reliability, and those marked with three or more are considered reliable. Since Croatia joined the European Union in 2013, the IATE database has tripled the number of Croatian terms from the existing 10,000. The number of Croatian terms today exceeds 36,000 and continues to grow rapidly. According to data obtained from the European Commission, Croatian terms in the IATE database have been validated as terms of high quality (over 80%) since 2015. They have been processed with added value in the past six years. The added value of a term means that in addition to the mandatory term and term reference in the target language, there is also a definition, definition reference, term context, context reference, term note, or language usage note.

In the past three years, European as well as global attention has focused on the terminology of the COVID-19 pandemic. Lexicography notes that social changes result in linguistic and terminological changes in the current global pandemic.

New terms emerged globally during the outbreak of COVID-19. The creation of new terms also led to the expansion of knowledge. Medical terms belonging to different fields have frequently been globally used in social media and press conferences. Linguists have analysed pandemic terminology from different points of view (cognitive linguistic, lexicographic, normativist, etc.), and coronavirus dictionaries and glossaries have been published in many countries, including Croatia (e. g. Coronavirus Glossary / Pojmovnik koronavirusa).

In this presentation, we will analyse Croatian pandemic terminology in the IATE database. Some of the pandemic terms belong to the medical field, e.g., names of viruses, strains, and drugs. This group also includes terms already defined by the ISO standard, e.g., medical equipment. Some terms refer to legislation (e.g., movement measures, restrictions, lockdown; this group also includes terms that have social significance, e.g., physical distancing and social distancing). Our analysis will start with concepts in the IATE database, which are processed primarily in the domain of SOCIAL QUESTIONS> Health and are marked with at least three asterisks (= reliable). The monolingual terms (in English) and bilingual terms (in English and Croatian) were extracted by the IATE extraction tool. Our analysis focuses on terms used differently in a specialised professional and legislative context and in the media (e.g., social distancing, lockdown, EU digital COVID certificate, recovery certificate, respiratory ventilator, artificial respirator, etc.). These terms are compared with terms in national terminology databases (e.g., STRUNA and Croatian Terminology Portal) and their equivalents in other official languages of the European Union recorded in relevant sources.

Keywords: Coronavirus terminology, COVID-19 terminology, IATE.

References

- Roche, Christophe; Alcina, Amparo; Costa, Rute. 2019. Terminological resources in the digital age. *Terminology* 25/2. 139–145.
- Temmerman, Rita. 2020. Terminology across languages: Researching neologisms and terminological variation in multilingual settings. *Svijet od riječi. Terminološki i leksikografski ogledi*. Eds. Ostroški Anić, Ana; Brač, Ivana. Institute of Croatian Language and Linguistics. Zagreb.
- Cabré Castellví, M. Teresa. 2003. Theories of terminology: Their description, prescription and explanation. *Terminology* 9/2. 163–199.

Štrkalj Despot, Kristina; Ostroški Anić, Ana. 2021. Rat protiv ratne metafore: metaforički okviri u hrvatskome diskursu o pandemiji koronavirusa. *Rasprave: Časopis Instituta za hrvatski jezik i jezikoslovlje* 47/1. 173–208.

Mihaljević, Milica; Hudeček, Lana; Lewis, Kristian. Corona-related neologisms – a challenge for Corona-related Neologisms – from Glossary of Coronavirus to Croatian Web Dictionary – Mrežnik.

Kristian Lewis¹

Marin Mrčela²

¹ Institute of Croatian Language and Linguistics, ² Supreme Court of the Republic of Croatia
klewis@ihjj.hr, marin.mrcela@vsrh.hr

Croatian legal terminology and Clear Writing movement: do citizens really understand judicial decisions?

In this talk, we shall pay attention to Croatian legal terminology in the context of the Clear Writing movement. Several fundamental presumptions need to be addressed when discussing this subject: a) language is a means of transmitting a legal norm and its application in a specific subject and in everyday life; b) legal norm and its application should be understandable to the recipients; c) clarity of judicial decisions depends on the person writing them. The Clear Writing movement provides certain guidelines that could be a contribution to the issues considered. Namely, it points out the problem of too long sentences containing many ideas, jargon words, and incorrect usage of terms. In addition, linguistic accuracy, clarity and comprehension have been detected as a problem in the process of compiling and editing legal texts. Having that in mind, we shall focus on particular issues such as the length of reasoning of the judgement, the usage of appropriate legal terms, and the introduction of the new (foreign) terms into Croatian legal terminology under the influence of the common EU legal framework. We believe that the proper usage of language in law requires the cooperation of legal experts and linguists.

Keywords: Clear Writing, legal language, legal terminology.

References

- Barbić, Jakša (ed.). 2013. *Jezik u pravu*. Hrvatska akademija znanosti i umjetnosti, Zagreb.
- Barbić, Jakša. 2013. Jezik u propisima. *Jezik u pravu*. Ed. Barbić, Jakša. 49–77.
- Lewis, Kristian. 2016. Četiri riječi, novi mačci u vreći – aplikacija, autoritet, divizija i rezolucija, *Hrvatski jezik* 3/3. 31–34.
- Lewis, Kristian. 2020. False Friends in Terminology: Croatian Lost in Translation, *Rasprave: Časopis Instituta za hrvatski jezik i jezikoslovlje* 46/1. 141–164.
- Mrčela, Marin. 2013. O jeziku u presudama. *Jezik u pravu*. Ed. Barbić, Jakša. 79–91.
- Mrčela, Marin; Lewis, Kristian. 2016. *Pravno nazivlje i hrvatski jezik*. Udruga hrvatskih sudaca. Zagreb.

Mrčela, Marin; Lewis, Kristian. 2017. I u pravu treba biti u pravu – o jeziku hrvatskih pravnih tekstova. *Hrvatski jezik* 4/3. 32–34.

Šarčević, Susan. 2013. Jezik kao izraz prava. *Jezik u pravu*. Ed. Barbić, Jakša. 21–35.

Željka Macan¹

Luka Medak²

¹ Faculty of Humanities and Social Sciences, University in Rijeka

² Faculty of Humanities and Social Sciences, University of Rijeka (graduate student)

zmacan@ffri.hr, luka.medak@live.com

Terminology in the Manual Official International Rules for Basketball

With new sports emerging in the Croatian speaking area, there was a need to publish manuals explaining the tactical and technical rules of each sport, which brought about formation of sports terminology and its usage. Croatian sports terminology was (and still is) strongly influenced by the English language, which will be confirmed by the examples of terms verified in the *Official International Rules for Basketball* manual.

The manual *Official International Rules for Basketball* was published by *Gimnastikon*, a professional journal for physical culture. The author and the year of publication were not mentioned; however, the Preface was signed by a pedagogue and sports worker Vladimir Janković, who came to Zagreb in 1932 after studying in Berlin and was the so-called skills (physical education) teacher in the First Men's Modern Grammar School. The manual also contains several advertisements for basketball equipment from 1933, so it can be assumed that it was published then.

The research aimed to determine and describe the similarities and differences between the terms confirmed in the source used and those that are in active use nowadays. In accordance with that, two hypotheses have been formulated. The first one is that the Manual will confirm the terms taken directly from the English language, which have been substituted by the domestic terms in today's terminology, which is based on terminological development, planning and standardisation. The second hypothesis is that some of the basketball terms confirmed in the Manual nowadays have the status of historicisms, i.e. they belong to passive vocabulary for extralinguistic factors, primarily due to changes in the rules of the game of basketball.

From the text of the manual *Official International Rules for Basketball*, 37 basketball terms are excerpted, analysed at a semantic level, and afterwards the degree of their adaptation to Croatian as the recipient language is being considered. The collected terms are divided into several groups, as follows: 1. Terms that are still in active use today, 2. Terms that alternate with another term (terms), coexist with it/them in active use, 3. Terms that are at the synchronic level replaced by new ones, and 4. Terms that became historicisms, that moved into a passive vocabulary. Within these subgroups, the collected terminological material is observed according to the criterion of

adapting certain terms to Croatian as the recipient language, followed by the division into foreign words, unadapted, adapted and adopted loanwords. A special subgroup are calques and complex lexical expressions, multi-word terms that can be considered as phrases of zero desemantization and part of phraseology in a broader sense.

In the presentation, selected examples will be singled out for each of the above mentioned groups. Each of the terms is accompanied by a definition and an example of its usage in the online and lexicographic sources used. For terms in which time stratification occurs, the corresponding term in active language use is also given in the appendix, again with an indication of its use in the context. Also, among the listed words are those terms in which the coexistence of several names in active use at the synchronic level has been confirmed and those are also explained with the appropriate examples of their usage in context in comparison to those confirmed in the source used for the research. In cases of coexistence of several terms confirmed in the manual *Official International Rules for Basketball* and in meaningfully matching terms in active use, they are observed according to the principles of terminology stated in the project Croatian Special Field Terminology – Struna. Terms in active use are taken from the Official Basketball Rules and the Official Basketball Rules authorised by the Association of Croatian Basketball Referees, compared to the official FIBA basketball rules in English.

Keywords: *Official International Rules for Basketball*, basketball terms, adaptation to Croatian, terminological development.

References

- Anić, Vladimir. 2000. *Rječnik hrvatskoga jezika*. Novi Liber. Zagreb.
- Čilaš, Marica. 2001. Sportsko nazivlje u hrvatskome i makedonskom standardnom jeziku na početku 21. stoljeća. *Rasprave* 27/1. 291–302.
- Ćurković, Dijana. 2013. Prevedenice u sportskoj terminologiji: košarka. *Croatica et Slavica Iadertina* 9/1. 63–81.
- Frančić, Anđela; Hudeček, Lana; Mihaljević, Milica. 2006. *Normativnost i višefunkcionalnost u hrvatskome standardnom jeziku*. Hrvatska sveučilišna naklada. Zagreb.
- Jardas Duvnjak, Ivana. 2014. Anglicizmi u sportskoj terminologiji u hrvatskom jeziku. *Zbornik radova Veleučilišta u Šibeniku*. Eds. Zlatović, Dragan. Šibenik. 185–194.
- Kapetanović, Amir. 2005. Historizmi i semantičke promjene. *Rasprave* 31/1. 153–163.
- Lukenda, Marko. 1984. Rad na normiranju terminologije. *Rasprave* 10–11/1. 81–86.
- Naklada Gimnastikona. 1933. *Službena međunarodna pravila za basket-ball*. Zagreb.
- Upravni odbor Udruge hrvatskih košarkaških sudaca. 2014. *Službena košarkaška pravila*.

Stručna komisija Udruge hrvatskih košarkaških sudaca. 2018. *Službena pravila košarkaške igre*.
FIBA. 2018. *Official basketball rules*.

Marko Markovikj¹

Jovana Dobрева¹

Mary Lucas²

Irena Vodenska³

Lou Chitkushev²

Dimitar Trajanov^{1,2}

¹ Faculty of Computer Science & Engineering, ss. Cyril and Methodius University – Skopje, MK

² Computer Sciences Department, Metropolitan College, Boston University – Boston, MA, US

³ Administrative Sciences Department, Metropolitan College, Boston University – Boston, MA, US

marko.markovikj@students.finki.ukim.mk, jovana.dobрева@finki.ukim.mk, mmlucas@bu.edu, vodenska@bu.edu, ltc@bu.edu, dimitar.trajanov@finki.ukim.mk

Terminology and topics analysis of tweets related to the COVID-19 pandemic

This research aims to analyse the terminology and topics that social networks users used to express their experiences with the COVID-19 pandemic. We used three datasets concerning different aspects of the COVID-19 pandemic to accomplish this goal. The first dataset⁹ is the smallest one and contains around 3,798 tweets regarding COVID-19 Onset, and it consists of tweets collected in March 2020 using the #Covid hashtag. The second dataset¹⁰ has more than 380,000 tweets regarding COVID Vaccination, collected using the #CovidVaccine hashtag in the period from January 2020 until March 2022. The third dataset is collected by our team, and its goal is to get the users' comments regarding Long COVID. We collected 1.5 million tweets from March 2020 until March 2022 that mention “Long COVID” or have the #LongCovid hashtag.

The analysis process consists of preprocessing, modelling, and visualisation (Provost 2013). In the first phase of our analysis, the goal was to clear the content of the tweets and convert them into a list of terms that will be input into the modelling phase (Camacho-Collados 2018). A tweet consists of 140 characters and, apart from text, can contain hashtags, mentions, URL links, and emoticons. A hashtag is a word or phrase preceded with "#," so we decided to keep them and convert them into terms. Mentions, emoticons, and links are not common words or topics, so we filtered them from the tweet content. We also filtered the special characters, so as a result of this

⁹ www.kaggle.com/datasets/datatattle/covid-19-nlp-text-classification

¹⁰ www.kaggle.com/datasets/kaushikuresh147/covidvaccine-tweets

phase, we have tweets that contain only regular text. In the next step, we apply "white space" tokenization to create a list of tokens. In the next step from this list, we removed the standard English stopwords and a list of custom manually-selected stopwords based on the COVID-19 corpus. At the end of this preprocessing phase, we converted each tweet to a list of terms.

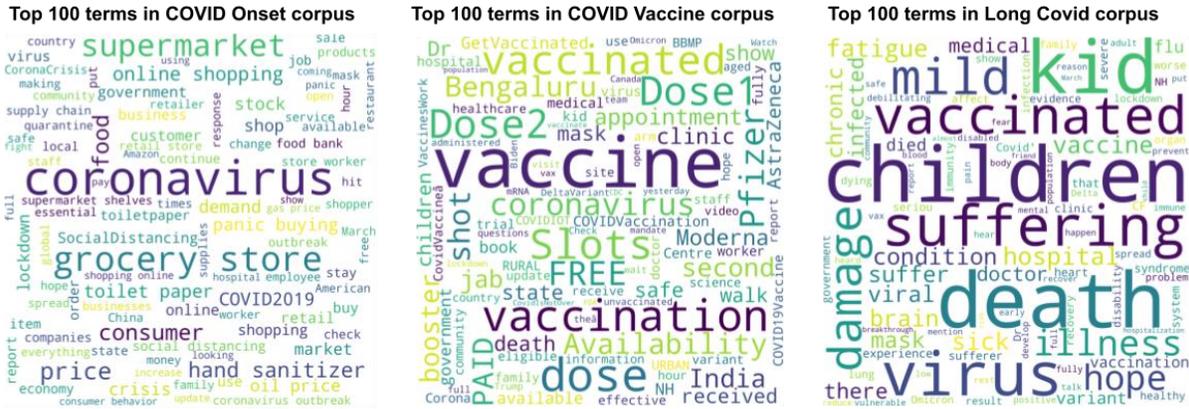


Figure 1: Word Clouds for Top 100 most frequent terms in the Onset Corpus, Vaccine Corpus, and Long COVID Corpus.

In the second modelling phase, we applied two different models. The first model uses the frequency of topics/terms to rank them to identify the top most used terms in each of the three different COVID datasets. To visualise the model results, we created word clouds, where the size of the word is proportional to their frequency in the corpus. Figure 1 shows the top 100 terms of each COVID dataset.

We use a characteristic words discovery algorithm based on the ScatterText library (Kessler 2017). The algorithm calculates the relative frequency of the words in the analysed corpus and compares them with the relative frequencies of the words in the standard English texts. The top 10 characteristic terms for each COVID dataset are presented in Table 1. When the COVID-19 pandemic started (the COVID Onset dataset), people were afraid that stores would be empty due to the ensuing panic and fear, which is reflected in the characteristic words in the Onset dataset. The Vaccines corpus has terms related to vaccination.

Interestingly, we can identify two new terms that have become very common in the everyday language: “fully vaccinated” and “boosted”. The third dataset, related to Long Covid, shows that people were separated into the categories of being vaccinated or unvaccinated. Hence, we introduced two new terms, 'vaxxed' and 'unvaxxed.' Interestingly, the term 'immunocompromised' has become one of the most frequently used terms in everyday language.

Table 1: Top 10 characteristic terms appearing in the COVID corpora

COVID Onset Corpus	COVID Vaccine Corpus	Long COVID Corpus
panic	doses	vaxxed
sanitizer	health	omicron
supermarket	jab	unvaxxed
toilet	Pfizer	immunocompromised
help	shot	micro clots
food	booster	symptoms
store	safe	risk
prices	fully	cases
oil	good	infection
crises	world	suffering

The goal of the third analysis was to analyse the temporal dynamic of the terms related to COVID. For this purpose, we utilised the Term association algorithm from the ScaterText library. The term association algorithms work on two corpora and create a scatter plot visualisation of the terms based on their association with the given corpora. We divided the Long Covid Tweets corpus into two parts based on the tweet's time: pre-vaccines with the tweets in the period [March 2020 to March 2021] and post-vaccines with the tweets in the period [April 2021 to March 2022]. The visualisation of the performed analysis is shown in Figure 2, where the terms more frequently associated with the first period are in the left upper corner, and the terms related to the second (post vaccine) corpus are placed in the right lower corner.

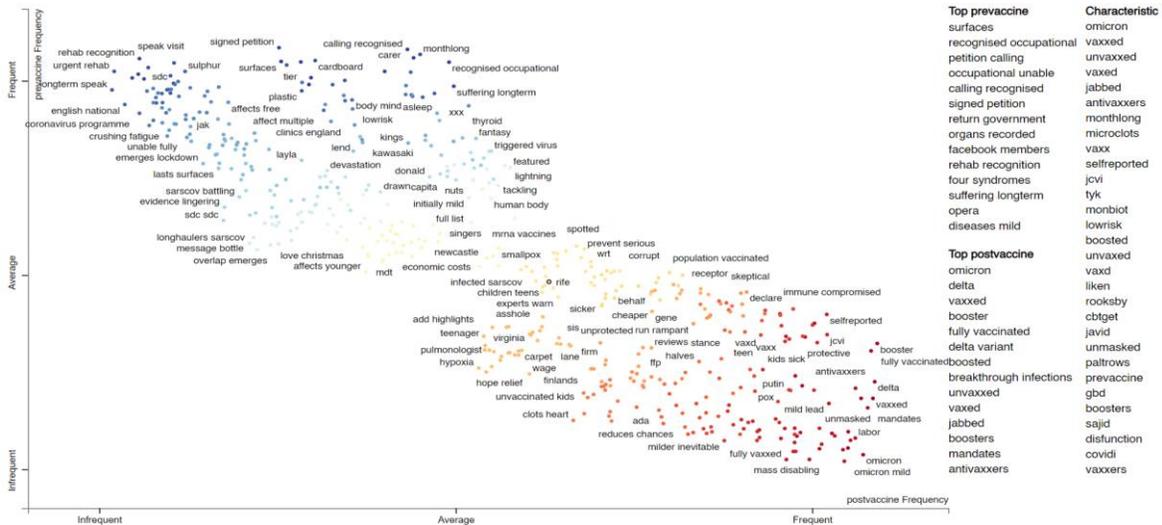


Figure 2: Scatter plot with terms association and the top characteristic terms for pre-vaccines and post-vaccines for Long COVID corpus

The presented analysis showed that the COVID-19 pandemic caused panic, especially from the existential perspective. At the onset, we observe fears of food shortages, high prices in supermarkets, or not having enough money for medications. Then, with the experience of Long COVID, the main concerns included lingering symptoms of COVID-19 infection and how they affect human health in the long term. With the development of Vaccines, discussions about advances in medicine surfaced and included vaccine technologies rapidly emerging due to the urgency of the pandemic.

We can conclude that human choice of language and terminology is time and context-dependent. The fluctuations in words and phrases used during a health crisis reflect the speakers' emotions and their search for and sharing of knowledge and experiences.

Keywords: terminology, COVID, long Covid, NLP, machine learning.

References

- Kessler, Jason S. 2017. *Scattertext: a Browser-Based Tool for Visualizing how Corpora Differ*. Association for Computational Linguistics. Vancouver, Canada.
- Provost, Foster; Fawcett, Tom. 2013. *Data Science for Business: What you need to know about data mining and data-analytic thinking*. O'Reilly Media, Inc.
- Camacho-Collados, Jose; Pilehvar, Mohammad T. 2018. On the role of text preprocessing in neural network architectures: An evaluation study on text categorization and sentiment analysis. *arXiv preprint arXiv:1707.01780*.

Ana Meštrović^{1,2}

Slobodan Beliga^{1,2}

Mihaela Matešić^{2,3}

¹ Faculty of Informatics and Digital Technologies, University of Rijeka

² Center for Artificial Intelligence and Cybersecurity, University of Rijeka

³ Department of Croatian Language and Literature, Faculty of Humanities and Social Sciences, University of Rijeka

amestrovic@uniri.hr, sbeliga@uniri.hr, mihaela.matesic@ffri.uniri.hr

Automatic Detection of Terminology and Topics from COVID-19 Related Texts Published on Social Media

The InfoCoV project* research aims to analyse online communication related to the COVID -19 pandemic. We applied natural language processing (NLP) and social network analysis (SNA) to the task of information monitoring, which is one of the four pillars of infodemic management (Eysenbach 2020). As a first step, we collected a large number of texts from online news portals and social networks. Next, we performed an extensive analysis of crisis communication using various NLP methods, such as keyword extraction, automatic extraction of the pandemic-related terminology, named entity recognition, topic modelling, thematic clustering, and sentiment analysis. We analysed how keywords, terminology, key entities, topics, and sentiment changed over time during the first three waves of the pandemic (Beliga et al. 2021). Our results show that despite the high number of COVID-19 related articles, key terms, and entities have encountered slow changes. In addition, we performed thematic clustering and sentiment classification of COVID-19 related texts posted on Twitter (Babić et al. 2021). We identified ten thematic clusters and tracked changes in themes and across three pandemic waves. According to our results, the majority of sentiment is negative and there are no significant differences across topics.

Acknowledgements

This work has been supported in part by the Croatian Science Foundation under the project IP-CORONA-04-2061, “Multilayer Framework for the Information Spreading Characterization in Social Media during the COVID-19 Crisis” (InfoCoV)

Keywords: Covid-19, keyword extraction, sentiment analysis, topic modelling.

References

- Eysenbach, Gunther. 2020. How to fight an infodemic: the four pillars of infodemic management. *Journal of medical Internet research* 22/6. E21820.
- Beliga, Slobodan; Martinčić-Ipšić, Sanda; Matešić, Mihaela; Petrijević Vuksanović, Irena; Meštrović, Ana. 2021. Infoveillance of the Croatian Online Media During the COVID-19 Pandemic: One-Year Longitudinal Study Using Natural Language Processing. *JMIR public health and surveillance* 7/12. E31540.
- Babić, Karlo; Petrović, Milan; Beliga, Slobodan; Martinčić-Ipšić, Sanda; Matešić, Mihaela; Meštrović, Ana. 2021. Characterisation of COVID-19-related tweets in the Croatian language: framework based on the Cro-CoV-cseBERT model. *Applied Sciences* 11/21. 10442.

Ana Ostroški Anić

Ivana Brač

Institute of Croatian Language and Linguistics

aostrosk@ihjj.hr, ibrac@ihjj.hr

The function of semantic roles in linking general and specialised lexical resources

Many lexical resources rely on semantic role labelling in determining event participants and describing semantic structures of a given language. Starting from different theoretical backgrounds, each database applies its own methods and principles, which results in several sets of semantic roles currently in use. There are approaches in which more general semantic roles are applied, e.g. LIRICS (Petukhova and Bunt 2008) or VerbNet (Kipper, Dang, and Palmer 2000) with roles that are not specific to particular verb classes, and approaches where verbs and other lexical units are defined according to specific semantic roles, of which FrameNet is best known (Baker 2017). Introducing semantic roles into terminological databases helps with building the conceptual structure of knowledge categories, as well as with finding equivalents in different languages. Terminological resources mostly follow frame semantics and FrameNet's methodology, e.g. Puertoterm (León Araúz, Reimerink, and Faber 2009), Ecolexicon (Faber and Buendía Castro 2014), DiCoInfo and DiCoEnviro (L'Homme 2012). Intertwining semantic with terminological or conceptual relations has highlighted the need for linking general and special knowledge in a clear and explicit manner. This idea has been well addressed in developing specialised resources based on FrameNet's methodology (e.g. L'Homme, Robichaud & Subirats 2020).

AirFrame (airframe.jezik.hr) is a specialised lexical database in which aviation terminology is defined in semantic frames, adapting FrameNet's methodology to terminology work. Basic processes, events, and entities characterising the field of aviation are defined in frames that consist of core and non-core frame elements (FEs) as semantic roles with a specific function in each frame, then lexical units or terms, and relations between frames. Whereas in traditionally organised terminological databases professional knowledge is defined in hierarchically ordered data categories, knowledge bases such as AirFrame allow for expanding the hierarchical structure and including chronological and associative relations, especially on the level of frame-to-frame relations.

This paper presents the structure of the AirFrame database, focusing on its differences and similarities with both general language and terminological resources. Frame element (FE) types or thematic semantic roles are defined in more detail, with examples of annotated sentences extracted from a specialised corpus compiled for the task. The function of FE types is to organise fine grained, frame specific roles into semantically connected groups. E.g. by using the FE type AGENT, we can mark specific frame elements PILOT, PILOT_IN_COMMAND, STUDENT_PILOT or AIRLINE as bearing the same thematic role of an agent in the sentence. On the other hand, we can also use the same FE, e.g. ALTITUDE, in different FE type positions, as INITIAL_LOCATION or FINAL_LOCATION, as in the sentence: ^{THEME}_{[FLIGHT_PATH} The trajectory] allows ^{THEME}_{[AIRCRAFT} the helicopter] to continue its FLIGHT ^{INITIAL_LOCATION}_{[ALTITUDE} from the height of the cruise] ^{FINAL_LOCATION}_{[ALTITUDE} to an altitude of 300 m (1000 ft) above the heliport]. The AirFrame database was designed with a view of defining a specialised domain, but at the same with the intention of including more general linguistic information, and developing a model that enables linking resources of different origin and purpose. Along with 20 aviation specific semantic frames, there are also 30 general frames in AirFrame at the moment, i.e. 30 frames defining FEs and lexical units of general Croatian language. Terminology defined according to this model could be used in developing domain ontologies or specific applications, but also as a useful teaching aid.

Keywords: terminology, aviation terminology, semantic frames, semantic roles, SRL.

Acknowledgements

This work has been fully supported by the Croatian Science Foundation under the project UIP-2017-05-7169.

References

- Baker, Collin F. 2017. FrameNet: Frame Semantic Annotation in Practice. *Handbook of Linguistic Annotation*. Eds. Ide, N.; Pustejovsky, J. Springer Science+Business Media Dordrecht.
- Faber, Pamela; Buendía Castro, Miriam. EcoLexicon. 2014. *Proceedings of the XVI EURALEX International Congress. The User in Focus*. Eds. Abel, Andrea; Vettori, Chiara; Ralli Natascia. EURALEX. Bolzano. 601–607.
- Kipper, Karin; Dang, Hoa Trang; Palmer, Martha. 2000. Class-Based Construction of a Verb Lexicon. *AAAI/IAAI*. 691–696.
- León Araúz, Pilar; Reimerink, Arianne; Faber, Pamela. 2009. Puertoterm & Macrocosta: A Frame-based Knowledge Base for the Environmental Domain. *Journal of Multicultural Communication* 1. 47–70.

- L'Homme, Marie-Claude; Robichaud, Benoit; Subirats, Carlos. 2020. Building Multilingual Specialized Resources Based on FrameNet: Application to the Field of the Environment. *International FrameNet Workshop 2020. Towards a Global, Multilingual FrameNet. Proceedings, Workshop of the Language Resources and Evaluation, LREC 2020*. Eds. Torrent, Tiago; Baker, Colin F.; Czulo, Oliver; Ohara, K; Petruck, Miriam R. L. 94–102.
- L'Homme, Marie-Claude. 2012. Adding syntactico-semantic information to specialized dictionaries: an application of the FrameNet methodology. *Lexicographica* 28. 233–252. doi.org/10.1515/lexi.2012-0012.

Martina Pavić

Institute of Croatian Language and Linguistics

mpavic@ihjj.hr

A corpus analysis of adjectives in different registers of Croatian medical terminology

Traditional terminological models often do not correspond to either linguistic or sociocultural reality (Cabr  1999) and using them it is difficult to describe adjectives, i.e. words that indicate the properties of objects and phenomena marked by other parts of speech, and relations between them. Contemporary terminological studies are focused primarily on nouns that name the basic referents in nomenclature. However, verbs, adverbs and adjectives that also participate in shaping the knowledge of relevant structures are neglected (Pitk nen-Heikkil  2015). The terminological importance of adjectives will be shown in this paper by researching their role in Croatian medical terminology within the framework of modern theories of terminology – communicative theory of terminology (Cabr  1999), sociocognitive theory of terminology (Temmerman 2000) and frame-based theory of terminology (Faber 2009) – according to which the specialised meaning of a lexical unit is realised in context. The contribution of adjectives to systematicity, precision and classification of terms into semantic groups will be emphasised, assuming that this part of speech contributes to the high structural organisation of medical terminology and plays an important role in the classification of the field. The development of medical science has prompted the expansion of medical terminology in which there are vaguely ambiguous complex terms, creating a sort of lexical barrier (Bowker and Hawkins 2006).

The analysis of adjectives in the medical field was conducted by compiling two corpora in the *Sketch Engine* tool – a scientific corpus consisting of relevant scientific journals in the medical field, such as *Medicus*, *Lije ni ki vjesnik*, *Infektolo ki glasnik*, and a popular corpus, which includes the most popular online portals dealing with topics in the field of medicine, such as *PLIVAzdravlje*, *Ordinacija.hr* and *Cybermed*. The aim was to identify similarities and differences in the use of adjectives in these two registers of communication. The analysis of each corpus was conducted in three phases: the extraction of single-word and multi-word adjective candidates for terms, the list of adjective-noun collocations, and the list of most frequent adjectives. Based on the obtained results, a comparative analysis of these two corpora was conducted.

The terminological unit is observed in the context of use, in which its meaning and place in the conceptual system are determined. A semi-automatic method of extraction was applied, which

meant that the automatic extraction of the most frequent adjectives was followed by an exhaustive research of both corpora. We started from the assumption that the termhood of an adjective is determined by the noun with which it appears, i.e. the context in which it is realised. The criteria for determining the termhood of an adjective are proposed, namely the relevance of its meaning in a given area, the existence of a definition and collocation potential.

The adjective candidates that were different in the scientific and in the popular corpus were singled out, and it was investigated whether internationalisms prevail in the scientific register in relation to the popular one. Furthermore, the difference in the collocation potential of adjectives was observed and it was investigated which adjectives have wider and in which narrower collocation potential. The analysis of collocations sought to determine the values of attributes by which members of each semantic class differ and confirm the assumption that, due to the adaptation of texts to a wider readership in the corpus of popular texts, adjectives have a narrower collocation potential. The analysis of collocations sought to determine the values of attributes by which members of each semantic class differ and confirm the assumption that, due to the adaptation of texts to a wider readership, adjectives have a narrower collocation potential in the corpus of popular texts. A comparison of the frequency of internationalism and adjectives of domestic origin in these two registers concluded how often the adjectives of domestic origin have been used. Finally, a semantic classification of adjectives in Croatian medical terminology has been proposed.

Keywords: adjective, medical terminology, scientific corpus, popular corpus, collocations.

Acknowledgements

This work has been fully supported by the Croatian Science Foundation under the project UIP-2017-05-7169.

References

- Bowker, Lynne; Hawkins, Shane. 2006. Variation in the organization of medical terms: Exploring some motivations for term choice. *Terminology* 12/1. 79–110.
- Cabré, Maria T. 1999. *Terminology: Theory, methods, and applications*. John Benjamins Publishing Company. Amsterdam – Philadelphia.
- Faber, Pamela. 2009. The cognitive shift in terminology and specialized translation. *MonTi. Monografias de Traducción e Interpretación* 1. 107–134.
- Pitkänen-Heikkilä, Kaarina. 2015. Adjectives as terms. *Terminology* 21/1. 76–101.
- Temmerman, Rita. 2000. *Towards New Ways of Terminology Description: The Sociocognitive Approach*. John Benjamins Publishing Company. Amsterdam – Philadelphia.

Sigita Rackevičienė¹

Liudmila Mockienė²

¹ Mykolas Romeris University, ² Mykolas Romeris University

sigita.rackeviciene@mruni.eu, liudmila@mruni.eu

Terminology Studies at Mykolas Romeris University

The aim of the paper is to present terminology studies at the Bachelor programme *English for Specific Purposes and the Second Foreign Language* at Mykolas Romeris University (MRU) in Lithuania. The paper focuses on the activities performed in and skills developed by the study courses that include terminology teaching or are devoted to terminology collection and management, as well as terminology research performed by MRU students.

The study programme includes four courses intended for the studies of English for specific purposes and communication in the fields of business, public administration, law, and finance. The courses are devoted to development of skills of content and linguistic analysis of field-specific texts, as well as oral and written communication in the fields. These learning outcomes depend heavily on the field-specific terminology comprehension and usage which are the focus of the courses and are achieved by various text analysis assignments and practical vocabulary acquisition tasks.

The study programme also includes a course devoted to terminology collection and management. During this course, students develop a terminographic project of a chosen domain. The project includes the following activities: getting acquainted with the chosen domain; building a domain specific English corpus using computer software (*Sketch Engine*); extracting terms with various software tools; searching for term definitions and categorising them according to the subdomains they represent; establishing term equivalents in two other languages; extracting examples of term usage in context from the ad hoc and other available corpora, and, finally, compiling a trilingual termbase using terminology management software (*Terminologue*). In addition, students perform quantitative analysis of the self-complied corpus structure and the selected terminology using computer software (*Sketch Engine* and *MS Excel*). The course develops a wide range of linguistic data management skills: developing an input dataset (corpus) and extracting terminological data and metadata from it and other sources, storing, organising, managing and analysing the extracted data, as well as compiling a terminology database. In addition, the work on terminographic projects allow students to acquire specialized knowledge of the chosen domains.

Students, who become interested in terminology studies, can choose terminology-related topics for their Term papers and Bachelor theses. The topics of papers encompass various aspects of terminology, such as terminology translation issues (equivalence in legal translation, metaphoric term translation, consistency of term translation, etc.), terminological concept relations and structures, term formation in different languages, term usage in media, etc.

Development of students' research skills is one of the major objectives of the study programme. Therefore, students are also involved in the research projects conducted by professors. One of the most recent examples of student involvement in research was participation of the study programme student Kamilė Dimaitė in the project *Bilingual Automatic Terminology Extraction (DVITAS)*, carried out by MRU together with researchers of Vytautas Magnus University (Lithuania) and included as a use case in the COST Action *European Network for Web-Centred Linguistic Data Science (Nexus Linguarum)*. The student contributed to the manual terminological annotation of a parallel cybersecurity corpus which is used as a training dataset for development of deep learning systems for bilingual automatic terminology extraction. The cooperation proved to be very successful. At the beginning of the assignment, detailed guidelines were given on the annotation scheme and the functions of the annotation software (*QuickTag*); during the whole assignment, supervision was provided. At the beginning, the student needed help, but gradually developed the skills to deal with complex cases and take decisions autonomously. In total, 1610 English and 1586 Lithuanian terms and proper names used in the EU Cybersecurity Act were annotated by the student. On the basis of the annotated dataset, the student carried out further research for her Bachelor Thesis, in which she dealt with semi-automatic lemmatization of the annotated English and Lithuanian terms using two lemmatizer (*spaCy*, *Lemuoklis*) and *MS Excel* for pre- and post-processing of the data, thematic classification of the annotated terms and analysis of their structural patterns in the English and Lithuanian languages. Participation in the project allowed the student to develop skills of terminological data annotation, processing and analysis, as well as skills of autonomous research, which are the ultimate goal of the studies.

The presented review of various types of terminology activities carried out at Mykolas Romeris University reveal the wide range of skills that can be developed through terminology-related courses and assignments. They encompass specialised communication, terminology extraction and management, terminology annotation and analysis and numerous other skills which are important not only for terminology work, but for the linguistic data management activities on the whole. The added value of any terminology-related work is acquisition of specialised knowledge of the studied domains.

Keywords: language for specific purposes, terminology studies, terminographic project, terminology annotation.

Acknowledgements

The project “Bilingual Automatic Terminology Extraction” is funded by the Research Council of Lithuania (LMTLT, agreement No. P-MIP-20-282). The project is also included as a use case in the COST action “European Network for Web-Centred Linguistic Data Science (Nexus Linguarum)” (CA18209).

References of the software mentioned in the abstract

Centre of Computational Linguistics, Vytautas Magnus University. *Lemuoklis*.
http://donelaitis.vdu.lt/main.php?id=4&nr=7_2.

Explosion. *spaCy* (version 3.2). 2016-2022. <https://spacy.io>.

Gaois research group, Dublin City University. *Terminologue*. 2019. The cloud-based instance
Terminologue.org: www.terminologue.org.

Lexical Computing. *Sketch Engine*. www.sketchengine.eu.

Microsoft Corporation. *Microsoft Excel*. <https://office.microsoft.com/excel>.

Rokas, Aivaras. *QuickTag* (version 1.1.0b). 2020.

Siniša Runjaić¹

Ana Šverko²

¹ Institute of Croatian Language and Linguistics, ² Institute of Art History
srunjaic@ihjj.hr, asverko@ipu.hr

From Initial Terminological Project to a Richly Illustrated Vocabulary: The Development of the Croatian Glossary of Classical Architecture

The *Glossary of Classical Architecture* (KLAS) was a project financed by the Croatian Science Foundation as part of the *Croatian Special Field Terminology* (STRUNA) programme. This paper aims to discuss how the conclusion of this project led to the development of a new one: the future Illustrated Glossary of Classical Architecture in Croatian.

The concept of classical architecture includes the architecture of ancient Greece and Rome. Later styles, including Renaissance, Baroque, Neoclassical, and several Neo-Historicist styles, relied on and built upon it in creative ways. It was, therefore, classical architecture – its vocabulary, architectural elements, principles, and typologies – that determined the development and understanding of architecture generally. As a result, there are numerous excellent glossaries of classical architecture in a variety of languages. Until recently, such a valuable terminological resource did not exist in Croatian, and as a result special field terminology in academic and professional publications was frequently imprecise and inconsistent. English terminology was frequently used uncritically and unsystematically, without considering whether the use of such terms was semantically appropriate. In particular, little effort was made to determine if these adopted terms were compatible with the rules relating to phonology, morphology, and word formation in the Croatian language.

The first part of this analysis, therefore, focuses on the all-encompassing classical architecture terminology project that was carried out through collaboration between art and architecture historians, conservators, and experts from the Institute of Croatian Language and Linguistics. Between 2018 and 2020, this project provided the groundwork for the development of more appropriate and precise terminology in the field of classical architecture. In June 2021, the glossary was finally evaluated and released to the public as part of the searchable national terminology database Struna (<http://struna.ihjj.hr/browse/?pid=48>). It thus became the first structured linguistic tool in the Croatian language for understanding and studying the field of architecture from the earliest periods to the present and is available to the broadest possible audience.

From the moment that the interdisciplinary team began working on it, the development of a traditional term base for classical architecture, and the organisation of the linguistic materials it contains, were considered only the necessary first step in the development of the glossary, a kind of intermediate stage between the initial idea of systematising the terminology and an illustrated glossary. Because the language of architecture is largely a visual language, while working on the terminology, the team also selected visual materials that related to the terms in question. In doing so, they remained constantly aware of the fact that Croatian classical monuments offered an excellent source for illustrating classical historical forms. Upon conclusion of the terminological project, therefore, the team immediately began working on the next phase, which resulted in the project *The First Illustrated Croatian Glossary of Classical Architecture, With Translations Into Four Languages*. The project has already been partly funded by the private Adris Foundation and is set to conclude in 2023 with the publication of the illustrated glossary.

The second part of the present analysis will focus on the project team's current methodologies for transforming the terminological entries relating to a particular architectural term into more detailed entries for a future vocabulary or lexicon. Moreover, it will discuss in more detail the refinement and systematisation of numerous terms and the scope of the work done on terms for a more specialised subset of users, with the understanding that in contrast to the publicly available term base, the illustrated glossary will be used more by specialists and students. On the editorial level, moreover, consultation of the best international publications relating to this topic forms an important part of the work, with the aim of selecting an appropriate academically trained illustrator and the best possible adaptation of the Croatian linguistic materials to the selected visual materials. Finally, it is important to note the international element that contributes to the quality of the project: the inclusion of a team of consultants from the Getty Research Institute in Los Angeles. Their involvement makes this project even more valuable, as it ensures that the Croatian glossary will be linked with the structured professional resources that form part of their Getty Vocabularies programme.

This paper will conclude with a discussion of how a field that was fairly unsystematic when it came to terminology, as the field of Croatian classical architecture terminology was, can be transformed through five years of interdisciplinary work into a well-rounded digital and print resource for future professionals, allowing them to continue working with confidence on national heritage in an international context.

Keywords: Croatian terminology, glossary of classical architecture, illustrated vocabulary.

References

- Bauer, Silvia Cerrella. 2015. Managing Terminology Projects: Concepts, Tools and Methods. *Handbook of Terminology, Volume 1*. Eds. Kockaert, Hendrik J.; Steurs, Frieda. D. John Benjamins. Amsterdam. 324–340.
- Brbora, Sanja. 2012. *Djelotvornost ilustracije u rječniku: slikovni rječnici – mogućnosti i ograničenja*. Doctoral dissertation. Faculty of Humanities and Social Sciences, University of Zagreb.
- Gangla, Lilian Atieno. 2001. *Pictorial Illustrations in Dictionaries*. Master thesis. University of Pretoria.
- Hammerly, Hector. 1974. Primary and Secondary Associations with Visual Aids as Semantic Conveyors. *International Review of Applied Linguistics in Language Teaching (IRAL)* 12/2. 119–125.
- Lew, Robert. 2010. Multimodal Lexicography: The Representation of Meaning in Electronic Dictionaries. *Lexikos* 20. 290–306.