



**Sveučilište u Zadru**  
Universitas Studiorum  
Jadertina | 1396 | 2002 |

---

# THE 17<sup>th</sup> NOOJ INTERNATIONAL CONFERENCE 2023

## Book of Abstracts

**Zadar, May 31 – June 2, 2023**

---

## Editors

---

**Linda Mijić**

Department of Classical Philology  
University of Zadar  
Zadar, Croatia

**Anita Bartulović**

Department of Classical Philology  
University of Zadar  
Zadar, Croatia

**Marijana Tomić**

Department of Information Sciences  
University of Zadar  
Zadar, Croatia

**Laura Grzunov**

Department of Information Sciences  
University of Zadar  
Zadar, Croatia

**Kristina Kocijan**

Department of Information and Communication Sciences  
Faculty of Humanities and Social Sciences  
University of Zagreb  
Zagreb, Croatia

**Max Silberztein**

Université de Franche-Comté  
Besançon, France

## Publisher

University of Zadar  
For the Publisher: Dijana Vican, rector

**ISBN 978-953-331-421-1**

# Contents

---

<b>INVITED TALKS .....</b>	<b>1</b>
<b>From Corpora and Corpus Tools to Dictionaries and Beyond: Infrastructure for Slovene</b>	
Iztok Kosem .....	2
<b>Using Semantic Hypergraphs to Represent Knowledge in Natural Language Text</b>	
Branko Žitko .....	4
<b>MORPHOLOGICAL &amp; LEXICAL RESOURCES .....</b>	<b>6</b>
<b>Quechua-Spanish and Spanish-Quechua Electronic Dictionaries of Verbs for NLP</b>	
Duran Maximiliano .....	7
<b>A Nooj Dictionary for Italian Light Verb Constructions</b>	
Alessia Nicola, Giocondo Cirillo .....	9
<b>Attenuative Collocations and Parametric Verbs in Old French and Old Spanish: A Contrastive Study</b>	
Xavier Blanco, Rafael García Pérez .....	11
<b>German <i>Selbst</i>-Compounds: A NooJ Grammar</b>	
Marco Angster .....	13
<b>Spelling Error Detection and Correction for Arabic Using NooJ</b>	
Rafik Kassmi, Samir Mbarki, Abdelaziz Mouloudi .....	15
<b>Towards a Linguistic Annotation of Arabic Legal Texts: A Multilingual Electronic Dictionary for Arabic</b>	
Khadija Ait ElFqih, Maria Pia di Buono, Johanna Monti .....	17
<b>The Construction of a Multilingual Legal Ontology</b>	
Ismahane Kourtin, Samir Mbarki .....	19
<b>Automatic Translation of Continuous and Fixed Arabic Frozen Expressions Using NooJ Platform</b>	
Asmaa Kourtin, Asmaa Amzali, Mohammed Mourchid, Abdelaziz Mouloudi, Samir Mbarki .....	21
<b>Recognition of Frozen Expressions in Belarusian NooJ Module</b>	
Yauheniya Zianouka, Mikita Suprushuk, David Latyshevich, Yuras Hetsevich .....	23

# Recognition of Frozen Expressions in Belarusian NooJ Module

---

Yauheniya Zianouka  
Mikita Suprushuk  
David Latyshevich  
Yuras Hetsevich

United Institute of Informatics Problems  
Minsk, the Republic of Belarus  
ssrlab221@gmail.com

## Abstract

In modern science, the question of the initial structural unit and perception of speech has no unambiguous solution because of various approaches and principles. However, it is known that speech has a syntagmatic nature and comprises lexical units that form syntagms. The author's delimitation and right intonation provide an adequate perception of speech. But synthesized speech, which is presented in various applications with voice accompaniment, is absorbed as unnatural, illegible, and inexpressive. The way to solve this problem is to develop specific methods and algorithms for analyzing and processing intonation features of natural speech, its automatic syntagmatic separation and implementation of all intonation constructions of a given language in NooJ. It will lead to automated reproduction of arbitrary text with the manner of human reading, and not an artificial system.

At the previous stages of our research, we have composed syntactic grammars for extracting syntagms at the punctuational and lexical levels, highlighting the intonation boundaries (Hetsevich et al., 2016; Zianouka et al., 2021). The next task is to form a syntactic grammar for delimiting phraseological units or so-called "Frozen expressions". These phenomena are reproducible, at least two-component linguistic units that combine with words of free use and are integral in meaning. As a rule, they are stable in their composition and structure (that is idioms). But in Belarusian, there are many frozen expressions that complicate their search due to structural units. For instance, the composition of phraseological expressions can be replaced by synonyms or other separate words. Or combinations, where one of the components is used in a phraseologically related meaning, and the other in a free one. Another problem is the order of units: in the phraseology, it can be fixed or, more often, it is used with the reverse order.

So, it is planned to compile a phraseological dictionary of the Belarusian language in NooJ format based on Etymological dictionary of Belarusian phraseological units and annotate it. The next step is to build syntactic grammars for searching the most typical groups of frozen expressions. And finally, to test them on the literary corpus of Belarusian NooJ module. This will contribute to the study of automatic processing of phraseological units for their further extraction into separate syntagms and forming their intonation portraits.

*Acknowledgements. The research is prepared under the project N<sup>o</sup>20213054 of The Belarusian Republican Foundation for Fundamental Research.*

## Key words

*Frozen expression, syntagma, intonation, delimitation, automatic processing*

## References

- [1] Hetsevich, Y., Okrut, T., Lobanov, B. (2016) Grammars for the Sentence into Phrase Segmentation: Punctuation Level. In: Okrut, T., Hetsevich, Y., Silberztein, M., Stanislavenka, H. (eds.) *Automatic Processing of Natural-Language Electronic Texts with NooJ*, 74–82. Springer, Cham.
- [2] NooJ: A Linguistic Development Environment, <http://www.nooj-association.org>, last accessed 2023/01/03.
- [3] Silberztein, M. (2016) *Formalizing Natural Languages: The NooJ Approach*. Wiley-ISTE, London.
- [4] Zianouka, Y., Hetsevich Y., Latyshevich, D., Dzenisiuk, D. (2021) Automatic Generation of Intonation Marks and Prosodic Segmentation for Belarusian NooJ Module. In: Bigey, M., Richton, A., Silberztein, M., Thomas, I. (eds.) *Formalizing Natural Languages: Applications to Natural Language Processing and Digital Humanities*, 231–242. Springer, Cham.