



C.R.I.T.



Centre de Recherches
Interdisciplinaires
et Transculturelles
EA 3224

16th INTERNATIONAL NOOJ 2022 CONFERENCE

Book of Abstracts

Hybrid Conference

Rosario, Santa Fe, Argentina

June 14-16, 2022

Conference Venue

**ECU (Espacio Cultural Universitario, UNR)
Rosario, Santa Fe, Argentina**

**Silvia Susana Reyes
Andrea Rodrigo
Max Silberztein
Carolina Tramallino (Eds.)**

Table of Contents/Sumario

RA: A New Linguistic Engine for NooJ Max Silberztein	5
Morphological & Lexical Resources	6
The architecture of Santi-Morf's Guesser Module Prihantoro	7
NooJ Morphological Grammar of Adjectives and Adverbs for Medieval Latin Linda Mijić & Anita Bartulović	8
Formalizing the Ancient Greek Participle Inflection with NooJ Silvia Susana Reyes	9
Corpus Analysis with NooJ: Mining of Spanish Neologisms Silvana Pierabella	10
Automatic analysis of appreciative morphology. The case of nominal paronomasia in Colombian Spanish Walter Koza, Viviana Román González & Constanza Suy	11
A computational approach to recognition of deverbal nouns ending in <sión> María José González	12
Syntactic & Semantic Resources	13
Ukrainian productive morphological grammars for recognizing unknown units Olena Saint-Joanis	14
Annotation of procedural questions in standard Arabic using syntactic grammars Essia Bessaies, Slim Mesfar & Henda ben Ghazela	15
Zellig S. Harris' Transfer Grammar and its application with NooJ Mario Monteleone	16
Formation and Evolution of Intensive Adverbs in <i>-mente</i> derived from the Adjectival Class <Causatives of Feeling> Rafael García Pérez & Xavier Blanco Escoda	17
Detection of adjunct prepositional phrases expressing time in the output of Estonian students of Spanish as third language: a computational approach Virginia Rapún Mombiela, Carolina Tramallino & Romina Arnal	18
Directed Motion Verbs in the Interlanguage of Spanish learners: Automatic Analysis using the NooJ system oriented to the Learning of Spanish as a Second and Foreign Language Romina Arnal, Carolina Tramallino & Virginia Rapún Mombiela	19
Approach to locative constructions: using NooJ to automatically analyze Spanish <i>donde</i> (where)-structures Silvina Lorena Palillo & Andrea Fernanda Rodrigo	20
Localization of passive voice structures in corpus of scientific research articles in Spanish using the NooJ system Carolina Paola Tramallino, Celina Beltrán & Romina Paola Arnal	21
Finding English Phrasal Verbs in Large Corpora Peter A. Machonis	22
Lexicon-grammar tables for discontinuous Arabic frozen expressions Asmaa Kourtin, Asmaa Amzali, Mohammed Mourchid, Abdelaziz Mouloudi & Samir Mbarki	23
Syntactic analysis of complex sentences containing Arabic psychological verbs in NooJ Platform Asmaa Amzali, Asmaa Kourtin, Mohammed Mourchid, Abdelaziz Mouloudi & Samir Mbarki	24
Formalization of transformations of complex sentences in Quechua Maximiliano Duran	25
Automatic Translation of Arabic Legal Terminology Using NooJ Khadija Ait ElFqih, Maria Pia di Buono & Johanna Monti	26
Corpus Linguistics & Discourse Analysis	27
Analyzing Political Discourse: Finding the Frames for Guilty and Responsible Krešimir Šojat & Kristina Kocijan	28
Creation of a legal domain corpus for the Belarusian NooJ module: texts, dictionaries, grammars Valerii Varanovich, Mikita Suprunchuk, Yauheniya Zianouka, Tsimafei Prakapenka & Anna Dolgova & Yuras Hetsevich	29
Prosodic segmentation of Belarusian texts in NooJ Yauheniya Zianouka, David Latyshevich, Yuras Hetsevich & Mikita Suprunchuk	30

Valerii Varanovich

Belarusian State University, Kurchatova str. 5, Minsk, Belarus

Mikita Suprunchuk

Minsk State Linguistic University, Zakharov str. 21, Minsk, Belarus

Yauheniya Zianouka

Tsimafei Prakapenka

United Institute of Informatics Problems, Surhanava str. 6, Minsk, Belarus

Anna Dolgova

Minsk State Linguistic University, Zakharov str. 21, Minsk, Belarus

Yuras Hetsevich

United Institute of Informatics Problems, Surhanava str. 6, Minsk, Belarus

ssrlab221@gmail.com

The current language situation in the Republic of Belarus is characterized primarily as state bilingualism. At the legislative level, two state languages are established – Belarusian and Russian. But despite the state bilingualism, the vast majority of legislative documents are implemented only in Russian. Thus, of the 26 codes of the Republic of Belarus, i.e. texts which are available on the National Legal Internet Portal pravo.by, 25 are officially adopted in Russian and only one in Belarusian. One of the main factors hindering the practical support of bilingualism in the legal sphere of the Republic of Belarus is the unresolved problem of high-quality and fast linguistic processing of large texts, which testifies to the relevance of high-quality machine translation. To handle the question of translating legislative documents into Belarusian, Speech Synthesis and Recognition Laboratory of UIIP NASB, in cooperation with specialists from Faculty of Social and Cultural Communications of BSU, have translated all codes of the Republic of Belarus into the Belarusian language using automatic services of corpus.by.

The next step is to collect all legislative codes of the Republic of Belarus in the Belarusian language in order to create a unified text corpus. This task is relevant for solving the following tasks. First of all, it is very important to perform primary analysis of legal domain corpus to find out the main linguistic peculiarities of this kind of corpus in comparison with Belarusian literary corpus. Secondly, we will be able to compose different types of dictionaries (Belarusian-Russian, Belarusian-English, Belarusian-English-Russian dictionaries). This question is very actual for Belarusian since there are very few translated Belarusian-foreign and foreign-Belarusian dictionaries of legal terminology – we are aware of 6 dictionaries that are different in their merits and significance for the ordering and development of Belarusian legal terminology. And the last – but not least – task is to develop special morphological and syntactical grammars for further prosodic analysis of legal texts. Automated syntagmatic delimitation is still not solved for the Belarusian language. That is why developing NooJ grammars will assist the process of creating a system of prosodic marks (including punctuation and intonation marks) and further automatic segmentation of Belarusian texts of the legal domain.

References

- Budějovice (2017), Czech Republic. Barone, L., Monteleone, M. & Silberztein, M. (Eds.). *Revised Selected Papers.*, June 9-11 (101-111).
- Computational platform for electronic text and speech processing corpus.by (2019). [Electronic source]. <https://corpus.by/>. Last accessed July 12 2018.
- Drahun A. (2019). Semi-Automatic Proofreading of Belarusian and English texts. Drahun, A., Hetsevich, Y., Bakunovich, A., Dzenisiuk, D. & Shynkevich, J. *International Conference NooJ 2019: Book of Abstracts*. Tunisia: Hammamet.
- Hetsevich Y. (2016). Semi-automatic Part-of-Speech Annotating for Belarusian Dictionaries Enrichment in NooJ. Hetsevich, Y., Varanovich, V., Kachan, E., Reentovich, I. & Lysy, S. *Automatic Processing of Natural-Language Electronic Texts with NooJ: 10th International Conference*, České.
- Hetsevich, M., Silberztein, H. *Stanislavenka, Springer International Publishing* (3-15).
- National Legal Internet Portal of the Republic of Belarus (2019). List of legal acts [Electronic source]. <https://pravo.by/document/?guid=3871&p0=H11800130>. Last accessed July 18 2019.
- Reentovich I. (2016). The First One-Million Corpus for the Belarusian NooJ Module. Reentovich, I., Hetsevich, Y., Voronovich, V., Kachan, E., Kozlovskaya, H., Tretyak, A. & Koshchanka, U. *Automatic Processing of Natural-Language Electronic Texts with NooJ: 9th International Conference, NooJ 2015*, Minsk, Belarus, June 11-13. Okrut, Y. *Revised Selected Papers*.