

THE 19TH NOOJ INTERNATIONAL CONFERENCE 2025

**Paris, June 11-13, 2025 | Institut national des langues et civilisations orientales
(INALCO)**

Celina Guadalupe Colussi
(CETEHPL, Facultad de Humanidades y Artes, UNR)
celinacolussi@hotmail.com

Recibido el 14 de agosto de 2025
Aceptado el 28 de agosto de 2025

Abstract

The 19th NooJ conference took place in the auditorium of the National Institute for Oriental Languages and Civilizations (INALCO) in Paris, France [2]. The conference was held on June 11, 12, and 13, 2025. Presentations were offered both in person and online via Zoom. Researchers from several countries participated in the event, including France, Italy, Belarus, Indonesia, Chile, Argentina, and Croatia, to name a few. The papers presented demonstrated an increase in linguistic diversity, in line with the ethos of the NooJ platform, which was designed by Max Silberstein [3]. The opening lecture, delivered by Christian Boitet (GETALP-LIG-UGA, Grenoble, France), also emphasised the scope of the approach. In his keynote speech, he raised the need to integrate natural language processing (NLP) with so-called large language models (LLMs). In other words, we need to create a sphere of true cooperation that allows us to build on the resources created by linguists and use LLMs for various tasks, such as pattern formation. The rest of the presentations were grouped around the central themes proposed for this conference: lexical, morphological, and semantic resources; local syntax; disambiguation; linguistic corpus; discourse analysis; and natural language processing applications.

Keywords: NooJ, NLP, cooperation, linguistic corpus, LLM.

19.ª CONFERENCIA INTERNACIONAL NOOJ 2025

**París, 11-13 de junio de 2025 | Instituto Nacional de Lenguas y Civilizaciones
Orientales (INALCO)**

Reseña del Evento

Resumen

La decimonovena conferencia de NooJ se celebró en el auditorio del Instituto Nacional de las Lenguas y Civilizaciones Orientales (INALCO), París, Francia [2]. Esta conferencia tuvo lugar los días 11, 12 y 13 de junio de 2025. Las ponencias se ofrecieron tanto en forma presencial como

en línea a través de Zoom. En este evento, participaron investigadores de varios países como Francia, Italia, Bielorrusia, Indonesia, Chile, Argentina, Croacia por citar a algunos de ellos. Los trabajos presentados mostraron un aumento en la diversidad lingüística, acorde con el espíritu que sustenta la plataforma NooJ, diseñada por Max Silberztein [3]. También se visibilizó una amplitud en el enfoque, propiciada por la conferencia de apertura, a cargo de Christian Boitet (GETALP-LIG-UGA, Grenoble, Francia). En dicha conferencia central, se planteó la necesidad de integración del procesamiento automático del lenguaje natural (PLN) con los llamados lenguajes grandes (LLM, Large Language Models). Es decir, la creación de un ámbito de verdadera cooperación que permita partir de los recursos creados desde la lingüística y aprovechar los LLM para diversas tareas como por ejemplo, la conformación de patrones. En cuanto al resto de las exposiciones, se agruparon en torno a los ejes centrales propuestos para esta conferencia: recursos léxicos, morfológicos y semánticos, sintaxis local, desambiguación, corpus lingüístico, análisis del discurso y aplicaciones del procesamiento del lenguaje natural.

Palabras clave: NooJ, PLN, cooperación, lingüística de corpus, LLM.

1. Introduction

The NooJ tool is a free software and a text processor created by Max Silberztein, in which one the automatic treatment of different natural languages can be made. NooJ has basic notions of the general properties of natural languages, and it allows the user to incorporate dictionaries and grammars to “teach” the automaton about these languages.

For a long time and particularly after the pandemia, the annual international conference has adopted a hybrid modality, facilitating the participation of a greater number of researchers. The event is held in different parts of the world and it has a duration of three days. For this reason, this year, the NooJ Conference took place in Paris, the previous year in Italy, and the year before that, in Croatia. The continuity of this proposal represents a great effort that the NooJ community makes every year to process the different languages in a common format and within a global vision.

Here, the Argentine team has been cooperating for about ten years with the NooJ project. Its focus could be defined as pedagogical and it aims at a concrete use of NooJ in the teaching and learning of languages. In this way, in 2022 Argentina became the host of the NooJ Conference giving a LatinAmerican view. The abstract written by Mariana and Carmen Gonzalez [1] showed the most outstanding aspects of this conference. Different works were presented connected to the structure and acquisition of the language, but also some works connected to discourse.

This year, the 19th NooJ Conference was hosted by Paris, not only the most widely spoken languages in the world like English, Spanish, French and German languages were present, but also the languages spoken by restricted areas like Persian, Ukrainian, Croatian, Belarussian, Indonesian, Romani, Arabic languages among others made their presentations at the congress. There were presentations in Latin and Quechua, too, that is highly important for the nurturing of the program. There even appeared a new term, the coinage of “NooJing”; this term was used by Ivana Zeger from Croatia to refer to the emergence of a linguistic approach for speech detection of Hope EDI, that is, in chunks of language referring to Equality, Diversity, and Inclusion.

The organizing committee was made up by the following people: Bénet Vincent (CREE, INALCO, France), Gadjeva Snejana (CREE, INALCO), Krylosova Svetlana, co-organizer (CREE, INALCO), Petković Divna, co-organizer (INALCO), Rabiei Marzieh (Université de Franche-Comté), Rabiet Victor (INALCO), Saint-Joanis Olena, organizer (CREE, INALCO), Silberztein Max, organizer (Université de Franche-Comté), Vassilioutchek Jeanna (CREE, INALCO) and Watabe Masako (Université de Franche-Comté).

The NooJ Conference was housed this year in INALCO (Institut National des Langues et Civilisations Orientales), a French centre of high teaching and investigation. There, classes about languages and cultures of countries outside Occidental Europe are taught and also studied.

2. Development

The program NooJ continues growing, it is being nourished in many parts of the world, each country giving its contribution. Now, we would like to show a brief summary of the event, showing the different activities in chronological order.

2.1. Wednesday, 11th June:

The opening session was in charge of Christian Boitet (Université Grenoble Alpes, France), and it was called: *Empirical and Linguistic approaches to NLP. LLM & TAL: la résistance serait-elle futile désormais? ≈ LLMs & NLP: Is resistance futile now?* This central conference gave clear ailments into the integration and cooperation of the investigations in NLP according to the different sources that are appearing with the coming of generative AI.

As regard the topic of **Lexical and Morphological Resources** there were the following presentations: Fabio Meroni (University of Bergamo, Italy) *Introducing the Mwotlap Module for NooJ*, Marzieh Rabiei (Université de Franche-Comté, France) *Linguistic and Morphological Analysis of Persian Verbs: Utilizing the NooJ Platform*, Olena Saint-Joanis (INALCO, France) *Formalizing Relative Adjectives and Participles with Morphological Grammars for Ukrainian, Affixoids in Action*, Kristina Kocijan, Krešimir Šojat (Faculty of Humanities and Social Sciences, Zagreb, Croatia), *A NooJ Journey Through Croatian Neoclassical Compounds, Updating the Belarusian NooJ Module* Valery Varanovich, Mikita Suprunchuk, Yauheniya Zianouka and this session finished with Yuras Hetsevich (Belarusian State University, Minsk, Belarus) *Integrating New Lexical and Grammatical Resources*.

Then, as regard the topic of **Local syntax and Disambiguation** there were also the following presentations: *The Annotation of Indonesian Clitics: a Comparative Analysis of three computational tools*, Prihantoro (Diponegoro University, Semarang, Indonesia), *Disambiguation in Medieval Latin*, Linda Mijić and Anita Bartulović (University of Zadar, Croatia), *Disambiguation Challenges and Solutions for Belarusian Texts in NooJ*, Yauheniya Zianouka, David Latyshevich, Yuras Hetsevich (United Institute of Informatics Problems, Minsk, Belarus), *Removing Ambiguities in Rromani Phrases*, Masako Watabe (Université de Franche-Comté, France), *Creation and Comparison of NooJ Grammars for Temporal Expressions in Ukrainian and Serbian*, Divna Petković (NYU, Paris), *Formalization Proposal for the Automatic Analysis of Verbal Idioms with Psychological Predicates in River Plate Spanish*, Walter Koza (Universidad de Chile).

Finally, the sessions had the vitamin booster with the **NooJ Tutorial: Developing Transformational Grammars with NooJ**, Max Silberztein.

The first day ended with a welcoming cocktail.

2.2. Thursday, 12th June:

This second day was the third session:

As regard the topic **Syntactic, Semantic and Morphological Resources**, different works were presented *Transformations and Paraphrases of Quechua Interrogative Sentences*, Maximiliano Duran (Université de Franche-Comté, France), *From Local Grammars to a Full Grammar: French Transitive Sentences*, Max Silberztein (Université de Franche-Comté, France) and finishing with *Media Framing of Migration: A Cross-Linguistic NLP-Based Semantic Analysis*, Maja Ivanović (Faculty of Humanities and Social Sciences, Zagreb, Croatia).

As regard the topic **Corpus Linguistics**, the talks were the following: *Unveiling the Language of Joy: A Computational Analysis of Happy Moment Expressions in NooJ*, Dijana Milek (Faculty of Humanities and Social Sciences, Zagreb, Croatia), *Mapping Tolstoy's Heart: Emotion Detection in Tolstoy's Works*, Lucija Bažant (Faculty of Humanities and Social Sciences, Zagreb, Croatia)

and *Emotion Recognition in Text*, Ivana Ujevi (Faculty of Humanities and Social Sciences, Zagreb, Croatia).

After this fourth session, lunch was taken. Afterwards, there was an excursion around the city of Paris, and the day finished with a Gala dinner.



Fig. 1 Sightseeing the city of Paris, on this occasion, visiting the Cathedral of Notre Dame

2.3. Friday, 13th June

This was the last day of the conference, and from Argentina, we contributed to the nurturing of the program. It was the fifth session:

In the topic, **Discourse Analysis and Natural Language Processing Applications**, the schedule was as follows: *NooJing for Hope: A Linguistic Approach to Hope Speech Detection*, Ivana Zeger (Faculty of Humanities and Social Sciences, Zagreb, Croatia), *Voices of Authority: A Study of Political Discourse and Ideology*, Lucija Bažant, Bruno Križić, Ana Košić (Faculty of Humanities and Social Sciences, Zagreb, Croatia), *Foreign Workers in the Spotlight: Linguistic and Sociocultural Insights from German and Croatian web portals*, Anja Brnjaković (Faculty of Humanities and Social Sciences, Zagreb, Croatia), *Automatic Processing of Expressions Using Numerals in a Student Corpus from the Universidad Nacional de Rosario (UNR)*, Carmen González, Celina Colussi, Iván Jesús Oliva (Universidad Nacional de Rosario, Argentina), *Use of LLMs for Semantic Relation Extraction and Development of NooJ Linguistic Software Applied to Drug Trafficking Language*, Nicolas Boffo (Université de Montpellier, France), *Development*

of a Didactic Web Interface in Ukrainian with NooJ, Victor Rabiet (ENS, Paris, France) and Processing the Language of Young University Students with NooJ, Andrea Rodrigo, Silvia Reyes and Mariana González (Universidad Nacional de Rosario, Argentina)¹.

In the topic, **Natural Language Processing Applications**, which opened with: *Emotional Weasels in Our Knowledge: A Tool to Depropagandise Discourse*, Tomislav Buki (Faculty of Humanities and Social Sciences, Zagreb, Croatia), *Detecting Grammatical Errors in Arabic using NooJ*, Rafik Kassmi (University Ibn Tofaïl, Morocco), and *Personalizing Learning Paths: Analyzing Student Interactions on Moodle Using NooJ platform*, Amira Abbes, H  la Fehri, Corinne Amel Zayani (Miracl, Sfax, Tunisia).

Then there was the closing of the NooJ 2025 Conference, and at the very end, they closed till next year with the General Assembly of the NooJ Association.



Fig. 2. Hybrid session about *Automatic Processing of Expressions Using Numerals in a Student Corpus* from the Universidad Nacional de Rosario (UNR)

3. Conclusion

The 19th NooJ Conference demonstrated the vitality of the NooJ community. Several presentations on discourse analysis stood out thanks to the enormous freshness of the proposals, establishing the topic as a particularly vital one. The presentations that stood out referred to: the detection of emotions in Tolstoy's texts, the media framing of migration, the treatment of semantic relationships in the language of drug trafficking², and an analysis including the term NooJing in the context of the so-called discourse of hope.

¹ In this part, we could show that we are also babysitting the program with the invaluable help and guidance of our godmother, Dr Andrea Rodrigo.

² This presentation: *Use of LLMs for Semantic Relation Extraction and Development of NooJ Linguistic Software Applied to Drug Trafficking Language*, by Nicol  s Boffo from the University of Montpellier, was very directly related to the opening conference.

Several presentations on disambiguation also stood out, including disambiguation in Belarusian texts, removing ambiguous phrases in Romani combining nouns and adjectives, and issues of disambiguation in medieval Latin. Important proposals were also presented, such as the updating of the Belarusian module with lexical and grammatical resources and the Mwotlap module within the NooJ software. In terms of the pedagogical application of NooJ, notable contributions included those analysis learning paths on the Moodle platform and the processing of Argentinian students' work, particularly with regard to numerals.

The conference was also enriched by a series of presentations focusing on: neoclassical compounds, prefixes, and suffixes derived from Latin and Greek in Croatian literature, relative adjectives and participles in Ukrainian, the approach to Indonesian clitics, temporal expressions in Ukrainian and Serbian, transformations and paraphrases of Quechua in interrogative sentences, the description of declarative sentences in French, and verbal idioms with psychological predicates in Rioplatense Spanish.

References

- [1] González, M., & González, C. (2023). XVI International NooJ Conference 2022: 14 al 16 de junio de 2022. Rosario- Argentina Reseña del evento. *Aprendo Con NooJ*, (3). <https://doi.org/10.35305/an.vi3.17>
- [2] NooJ Conference 2025. Available at: [NooJ 2025 International Conference - Sciencesconf.org](https://www.sciencesconf.org/conference/nj2025), [last consultation, August 28, 2025]
- [3] Silberztein, M. (2016): Formalizing Natural Languages:The NooJ Approach. Iste Ediciones, London.