LISBON SUMMER SCHOOL IN LINGUISTICS

PhD Programme

Lisboa, June 30 – July 4, 2025

Venue
Faculdade de Ciências Sociais e Humanas – NOVA
Avenida de Berna, 26-C
1069-061 Lisboa – Portugal

AREA 2: TERMINOLOGY AND LEXICOGRAPHY

Course 1: **The *Lexical Priming Theory* – how a new theory of words and language provided an explanation for corpus-based observations** - Michael Pace-Sigge (University of Eastern Finland)

9:30am - 12:30pm

COURSE DESCRIPTION

There must be a reason why words (or clusters of words) collocate, group into certain constructions but avoid other patterns of being meaning-giving and have certain semantic associations. Corpus Linguistics has been able to show us relations between words.

The Lexical Priming Theory (LPT), backed by psychological & cognitive research, gives us an idea why it is likely (or unlikely) that certain words (clusters of words) stand in such relations. The basic premise is that by repeatedly encountering a word in a particular context, an individual's knowledge of the word is accumulated; "as social beings and as an integral part of all our animate and inanimate surroundings, we are touched, influenced, and formed by what we are exposed to".

Providing an overview of Michael Hoey's (2005) theory, this seminar now aims through practical applications to show in how far the LPT is supported and can be employed for a variety of linguistic and lexical studies.

SHORT BIBLIO

Hoey, M. (various 1991-217)
Hoey, M. (2005) Lexical Priming. Routledge
Pace-Sigge, M. (2013) Lexical Priming in Spoken English. Palgrave Macmillan
Pace-Sigge M & Patterson K. eds (2017) Lexical Priming: Applications and Advances. John Benjamins.

BIO

https://uefconnect.uef.fi/en/michael.pace-sigge/

Course 2: **Linguistic Linked Open Data approaches for Lexicography, terminology and text corpora** – David Lindemann (UPV/EHU, University of the Basque Country)

2:pm - 5:00pm

COURSE DESCRIPTION

Day 1: Five Star Open Data, Linked Open Data, Semantic Web. A history of data representation formats, and requirements for lexicographical and terminological datasets and digital corpora: Towards FAIR (findable, accessible, interoperable, reusable) data, actionable by humans and by machines.
Day 2: Participants' background and own miniproject presentations. / Free software solutions for FAIR lexical datasets: Overview. XML-based solutions, RDF-based solutions. Inception, a tool for collaborative semi-automatic corpus annotation.
Day 3: Introduction to Wikidata, a free and collaboratively edited Linked Open Data resource, combining ontological and linguistic knowledge.
Day 4: Build your own Knowledge Graph using Wikibase. Use cases and applications. Collaborative work on participants' miniprojects.
Day 5: Wrap-up and discussion, final miniproject presentations. While it remains optional, participants are encouraged to bring own datasets for use in experiments, or even propose a miniproject to collaboratively work on. This can be any kind of lexical dataset (a wordlist, a text to annotate, a dictionary or term collection, etc.).

REFERENCES

. Five Star Open Data: https://5stardata.inf
. FAIR data: https://www.go-fair.org
. Linguistic Linked Open Dat
. Introduction and Overview (Chiarcos et al. 2013)
. Recent developments (Declerck et al. 2020)
. Wikidata and Wikibas
. FAIR data on Wiki-Platforms (Lindemann 2025)
. INCEpTION (Eckart de Castilho et al. 2018

Course 3: **Structuring Unstructured Data for Effective Knowledge Organization** – Margarida Ramos (NOVA FCSH | CLUNL) – ***limited to 15 attendees***

5:30pm - 8:30pm

COURSE DESCRIPTION

In the context of information systems, an ontology is a form of knowledge representation that models both real and imaginary objects—for instance, when defining the concept of a unicorn. From a computer science perspective, ontologies define sets of representational primitives used to model a domain of knowledge or discourse (Gruber, 2008). These primitives include classes, attributes (properties), and relationships between classes. Ontologies are widely used to structure knowledge for specific applications and domains, facilitating the development of knowledge-based systems that enable knowledge sharing and reuse. The process of modeling an ontology can be expressed through several types of definitions, such as mathematical, logical, or other structural languages. In this course, we will focus on the logical definition.

Ontologies typically represent distinct levels of abstraction, ranging from top-level ontologies to domain-specific ones. They serve as gateways to conceptualizations by defining vocabularies—such as terms denoting concepts—formally expressed through logical constructs.

This session will explore knowledge organization through the dual dimension of terminology: linguistic and conceptual, with a particular emphasis on the perspective of Digital Humanities. To semi-formally and formally represent knowledge, we will begin by analysing natural language texts to capture the vocabulary of a given field of interest—including terms denoting concepts, domain relationships, and the terminology used to express conceptualizations within that domain.

For knowledge representation, we will use different tools based on the degree of formality, distinguishing between semi-formal and formal language specifications. Protégé Desktop v.5.6.5 will serve as the ontology editor for formally modeling terminological data using the OWL 2 Web Ontology Language (Manchester Syntax), enabling the creation of a structured and logically coherent knowledge organization.

By the end of this session, participants will be able to (1) analyse a textual corpus, (2) manually represent a set of concepts in a semi-formal manner and (3) formally represent them using logical constructs. The first and second steps of this workflow will establish the foundational structure necessary to ensure the coherence of the third step, aligning with the ISO/TC 37 standards for conceptual relationships.

REFERENCES

Gruber, T. (2009). Ontology. In L. Liu, & T. Özsu (Eds.), Encyclopedia of Database Systems (pp. 1963-1965). Boston, MA: Springer. doi:https://doi.org/10.1007/978-0-387-39940-9_1318

Horridge, M. (2011). A Practical Guide To Building OWL Ontologies Using Protege 4 and CO-ODE Tools.

ISO/FDIS 1087 (E). (2019). Terminology work and terminology science - Vocabulary. Suisse: ISO.

Lim, E., Liu, J., & Lee, R. (2011). Knowledge Seeker – Ontology Modelling for Information Search and Management (Series: Intelligent Systems Reference Library ed.). (J. Kacprzyk, Jain, & Lakhmi, Eds.) Hong Kong: Springer Berlin, Heidelberg. doi:https://doi.org/10.1007/978-3-642-17916-7

Uschold, M., & Gruninger, M. (1996). Ontologies: principles, methods and applications. In The Knowledge Engineering Review (Vol. 11, pp. 93-136). Edinburgh: University of Edinburgh.

RESUMED BIOGRAPHY

Margarida Ramos: https://www.cienciavitae.pt/portal/4F1A-63BA-E128

Limited to 15 attendees

FEES

Fees for PhD Students at NOVA FCSH

- Enrolled in Doctoral Programs in Linguistics, Translation and Terminology, or Languages Teaching – Multilingualism and Education for a Global Citizenship: FREE
- Enrolled in any Doctoral Program at Universidade NOVA de Lisboa: 1st course. FREE; other courses: 50€ each

Fees for all

- Registration in a single course: 90€
- Registration in 2 courses: 120€
- Registration in 3 courses: 150€

NOTE: Special payment conditions for participants covered by protocols with NOVA FCSH (to be defined on a case-by-case basis, depending on the agreement).

REGISTRATION

Enrolment is open until 20th of June, 2025.

To enrol, please fill in the form available **here** and consider the following payment information.

Please, be aware that enrolment will only be validated after you send us the confirmation of transaction.

PAYMENT

Payment can be made by one of the following ways:

Option 1 – Bank transfer:

IBAN: PT50 0018 000321419114020 13

Banco Santander Totta S.A.

BIC/SWIFT: TOTAPTPL

Holder: FCSH – Univ. Nova de Lisboa

CIF / VAT number: 501559094

Address: Av. Berna 26 C, Lisboa 1069-061

Option 2 – Paypal:

Send payment to the account: dgfc@fcsh.unl.pt

In the transfer details, you must specify "Summer School – NOVA CLUNL".

Afterwards, please send the confirmation of transaction to sec.clunl@fcsh.unl.pt

ACCOMMODATION

You can find a list of nearby hotels **here**.