

Un'ontologia per la classificazione semantica dei dizionari storici: l'esempio del *LEI Digitale*



universität
wien

LA MEMORIA DIGITALE
XII CONVEGNO ANNUALE AIUCD
Università degli Studi di Siena,
5-7 giugno 2023

Alessandro Alfredo Nannini - University of Vienna

Keywords

Lessico Etimologico Italiano, Historical Linguistics, Onomasiology, Semantics, Lexicography, Digital Humanities, Italian Language, Conceptual Systems, Ontology

Summary

Si presentano qui le possibilità di riutilizzo, nell'ambito della lessicografia storico-etimologica italo-romanza, di un sistema concettuale costruito a partire dalla classificazione semantica del lessico anglosassone. Infatti, il progetto di mappatura semantica del *Lessico Etimologico Italiano* (LEI) utilizza come base noemica la tassonomia dell'*Historical Thesaurus of English*, opportunamente convertita in ontologia OWL (*Ontology Web Language*) e ampliata in base alle esigenze di classificazione del lessico trattato nel LEI. I risultati ottenuti fanno ben sperare anche nella prospettiva di un utilizzo comune da parte dei diversi progetti di lessicografia storica.

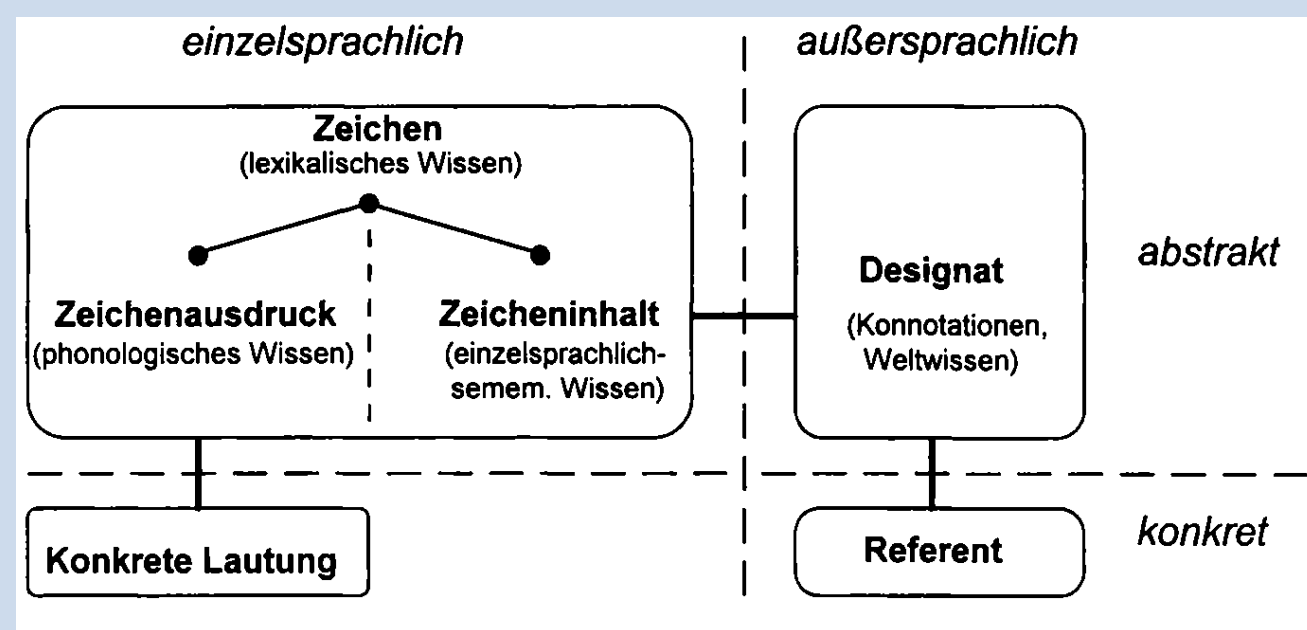
Theory

Lexical Semantics: Semasiology and Onomasiology

«**Semasiology** [...] considers the isolated word and the way its meanings are manifested» (Baldinger 1980, 278).

«The goal in **onomasiology** is to find the linguistic forms, or the words, that can stand for a given concept/idea/object» (Grzega/Schöner 2007, 7).

Onomasiology originated in the last years of the 19th century as a branch of Romance linguistics: the scholars investigated the different designations of a concept or a set of related concepts in the Romance varieties.



(Semiotic model by Blank 1997)

Theoretical principles

- Everything that is verbalized in a language is a concept;
- Onomasiological work builds upon semasiological work;
- Onomasiological work requires a conceptual network of reference that should not be seen as a pre-established universal system;
- This conceptual network is developed through semasiological analysis.

Proposals for further development

- Transformation of LEI data into RDF (*Resource Description Framework*) and modelling of lexical information using OntoLex-Lemon (see Tittel/Chiarcos 2018).
- Adoption of the same conceptual system by other humanities projects (especially historical lexicography).



<https://lei-digitale.it>

Motivation

Even though the first studies of onomasiology originated in the context of Romance languages, it is important to acknowledge that there is a **lack of onomasiological dictionaries** in the field of Romance lexicography, particularly in historical lexicography, whose typical scenario is characterised by semasiological dictionaries which «hamper[...] the semantic driven research of the *meaning* of the words» (Tittel et al. 2020, 1). Change of meaning and change of designation are fundamental phenomena in history of languages and cultures, and tools are needed to investigate and better understand these phenomena.

Main Aim

The main aim of the project is to develop a model for the **semantic mapping** of the largest historical dictionary of Italian language and its varieties, the *Lessico Etimologico Italiano* (LEI), in its new digital version (*LEI-Digitale*), in order to enable the semantic driven research (**from concepts to words**) within this extensive collection of lexical material.

Main Principles

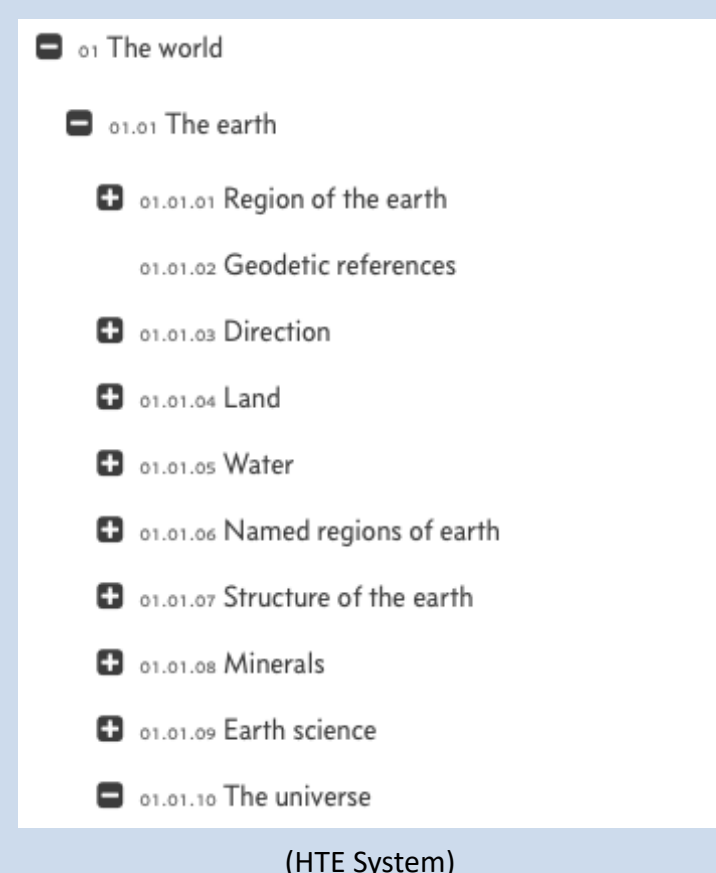
- The development of the conceptual system must **re-use** conceptual models that have been already used in lexicography; these models will be modified to meet the specific requirements of the LEI (see Theory).
- Use of modern and efficient technologies in the field of **e-lexicography**. The ongoing **digitisation** of the LEI and the resulting availability of data in electronic format enable it.

Work Steps

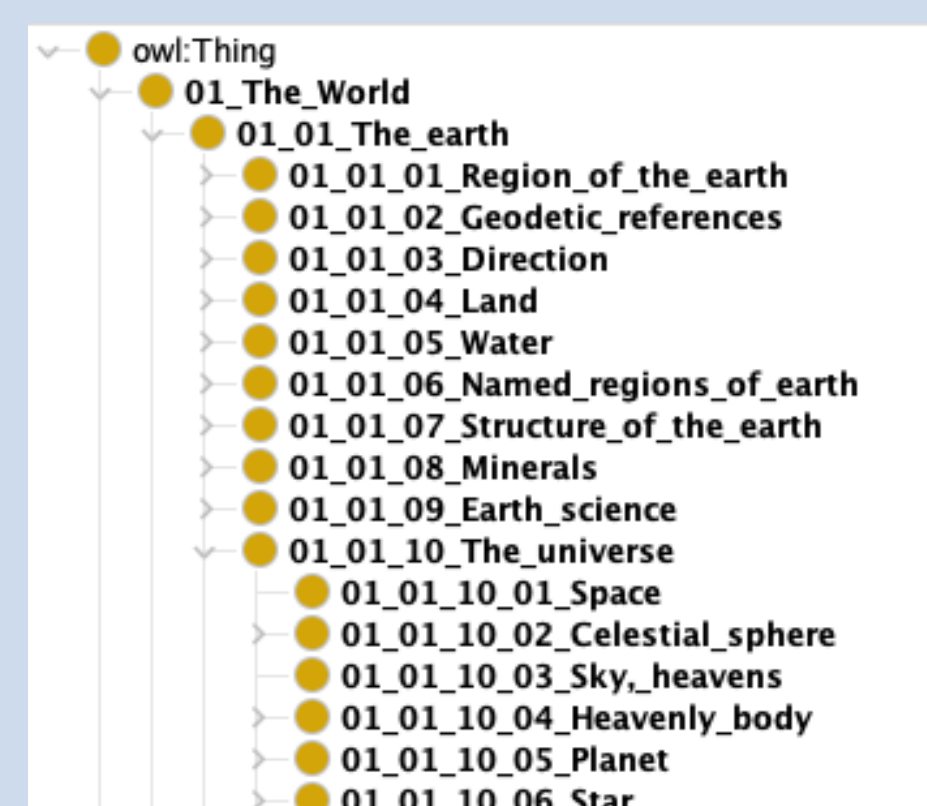
- Choose of the conceptual systems.** After a detailed analysis of conceptual systems used in European lexicography (both traditional and digital such Hallig-Wartburg and HTE), we decided to base the classification system of the LEI on that of **Historical Thesaurus of English** (HTE). Semantic mapping of some LEI articles demonstrated the effectiveness of the system.

- Development of LEI's conceptual system based on HTE's system**

→ Hierarchical data modelled as an ontology in OWL.



(HTE System)



(LEI Ontology)

- Semantic mapping stages:**

- Mapping of etyma at the macrostructural level;
- Mapping of semantic sections within articles;
- Mapping of the single components of the articles (i.e. The smallest sub-section of an article consisting of one or more lexical units having the same semantic features);

Mapping consists of adding conceptual information to the corresponding element in an XML document using an identifier via the *@corresp* attribute. During these stages, the initial system undergoes changes due to semantic characteristics of the lexicon;

- Embedding the new tool into the *LEI Digitale* web page;**

- Integration of the conceptual system into the LEI's new digital compiling system.**

References

- Baldinger, Kurt (1980): *Semantic theory. Towards a Modern Semantics*, Oxford, Basil Blackwell.
- Blank, Andreas (1997): *Prinzipien des lexikalischen Bedeutungswandels am Beispiel der romanischen Sprachen*. Tübingen, Niemeyer.
- Grzega, Joachim/Schöner, Marion (2007): *English and General Historical Lexicology: Materials for Onomasiology Seminars*. Eichstätt, Katholische Universität Eichstätt-Ingolstadt (Onomasiology Online Monographs), <http://edoc.ku-eichstaett.de/9521/> [24/05/2023].
- Hallig, Rudolf/von Wartburg, Walther (?1963): *Begriffssystem als Grundlage für die Lexikographie, Versuch eines Ordnungsschemas*, Berlin, Akademie-Verlag.
- HTE = Christian, Kay et al. (eds.). (2022): *The Historical Thesaurus of English* (2nd edn., version 5.0). University of Glasgow, <https://ht.ac.uk/> [13/01/2022].
- LEI = Prifti, Elton/Schweickard, Wolfgang (eds.). (1979-): *LEI. Lessico Etimologico Italiano*. Fondato da Max Pfister, Wiesbaden, Reichert; *LEI Digitale*: <https://lei-digitale.it> [24/05/2023].
- Tittel, Sabine/Chiarcos, Christian (2018): "Historical Lexicography of Old French and Linked Open Data: Transforming the resources of the Dictionnaire étymologique de l'ancien français with OntoLex-Lemon", in: *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC-2018)*, a cura di Nicoletta Calzolari et al., 58–66. Miyazaki, Japan: European Languages Resources Association (ELRA).
- Tittel, Sabine/Gilles-Webber, Frances/Nannini, Alessandro Alfredo (2020): "Towards an ontology based on Hallig-Wartburg's Begriffssystem for Historical Linguistic Linked Data", in: *Proceedings of the 7th Workshop on Linked Data in Linguistics (LDL-2020)*, Language Resources and Evaluation Conference (LREC 2020), Marseille, 11.–16. Mai 2020.



Alessandro Alfredo Nannini

Institut für Romanistik, Universität Wien
alessandro.alfredo.nannini@univie.ac.at