

A Model for Representing Diachronic Terminologies: the Saussure Case Study

Silvia Piccini <silvia_dot_piccini_at_ilc_dot_cnr_dot_it>, Institute for Computational Linguistics

Andrea Bellandi <andrea_dot_bellandi_at_ilc_dot_cnr_dot_it>, Institute for Computational Linguistics

Emiliano Giovannetti <emiliano_dot_giovannetti_at_ilc_dot_cnr_dot_it>, Institute for Computational Linguistics

Abstract

The aim of this article is to present a model for representing in an explicit and formal way the diachronic evolution of concepts and terms in a given domain, so that this formalization can be machine-actionable. The approach we here propose is based on Semantic Web technologies in order to guarantee interoperability and reuse across scientific communities of diachronic terminological resources that can be thus easily accessed, interconnected and mutually enriched. More specifically, the representation of dynamic evolution of terms and concepts was performed in OWL using the N-ary relations mechanisms. In addition, a set of SWRL rules was set up, in order to automatically identify the evolution of the concepts evoked within a text, as well as the terms representing these concepts. Our model was adopted to formally represent diachronic aspects of Saussure's terminology as they emerge from his works. An example will be provided to highlight the potential of such a knowledge structuration for gaining a wider understanding of the profound terminological and conceptual changes brought about by the paradigmatic and epistemological revolutions in sciences.

1. Introduction

Πάντα ῥεῖ. Everything is subjected to the inexorable law of change: the reality, the categories through which we organize it and the words we use to talk about it. 1

This is particularly the case for the history of science. Over the centuries, scholars have built different theoretical models in response to the continuous innovation that emerged from observation sometimes producing a real scientific revolution in the worldview. Needless to say, a change in conceptual level often corresponds to a change in terminology^[1]: new terms can be introduced to express the new system of concepts or old terms can be dismissed when the concept becomes obsolete. The aim of this article is to present a model for formally representing the diachronic evolution of concepts and terms in a given domain, so that this formalization can be machine-actionable. Taking advantage from past experiences conducted in the context of several projects^[2], we propose here an approach based on Semantic Web technologies that guarantees interoperability and reuse of diachronic terminological resources across scientific communities. 2

The development of standards for representing and exchanging data has always been a concern in the terminological field. Interoperability across terminological formats is basically ensured by adopting the ISO standard 30042: 2008 TBX (TermBase eXchange), an XML-based family of terminology exchange formats compliant with the Terminological Markup Framework (TMF - ISO 16642: 2003). 3

Over the last few years, interesting solutions for interoperability have been offered by the Semantic Web technologies and the Linked Open Data initiative which allow data to be shared, reused across applications and linked with the resources available on the Web, thus benefiting from the whole datacloud of URIs with which reference resources like 4

DBpedia are already networked.

In terminology, there has been a growing trend towards representing and publishing terminological resources on the Semantic Web according to Linked Open Data principles, as demonstrated for example by the NCBO (National Center for Biomedical Ontology) BioPortal, a repository of more than 300 biomedical ontologies and terminologies converted into RDF [Salvadores et al. 2013]; or by the Open Biomedical Ontologies (OBO) consortium [Smith et al. 2007] aiming to create a family of new or extant ontologies based on shared principles compliant with the Semantic Web technologies. These and other similar initiatives are all driven by the same need, namely to avoid that terminologies, created by proprietary software, exist independently of each other, hidden in isolated databases. The lack of semantic interoperability, indeed, precludes any possibility of progress, as medical data can be exploited to their full potential only if they are interconnected to each other. Sharing large quantities of data structured in ontologies makes it possible to use IT applications and artificial intelligence systems, enabling thus better diagnostics, early disease prevention and, more generally, a deeper understanding of the human biology and health.

Another initiative deserving to be mentioned is the EuroVoc management based on technologies in line with W3C recommendations and with the latest developments in classification standards. A collaborative project started in 2019 and undertaken by the Publications Office of the European Union and the International Labour Office [Dechandon et al. 2019] aims to align EuroVoc with other controlled vocabularies and publish them as Linked Open Data. The objective is to improve data accessibility and guarantee large-scale interoperability, thus facilitating the authoring-translating-publishing chain as well as the work in public administrations and international organisations.

Finally, Cimiano et al. (2015) proposed an approach to convert terminological datasets originally stored in TBX format to RDF by means of an online service named TBX2RDF, allowing thereby terminological resources to be used and enhanced across different scientific communities and, in particular, to become part of the great Linguistic Linked (Open) Data cloud (see *infra*).

Nevertheless, as we shall see, using these technologies to represent the diachronic evolution of terminological and conceptual data constitutes an interesting challenge, since OWL (Web Ontology Language), the W3C standard language developed to represent ontologies in the Semantic Web, is fundamentally static and monotonic.

2. Terminology and Diachrony

In the terminological literature, the diachronic dimension has been neglected for a long time. This was mainly due to the onomasiological, prescriptive and static approach underlying the General Theory of Terminology (GTT) elaborated by Wüster in 1968.

Wüster's approach to terminology was based on a neo-positivist referential and taxonomic semiotics [Slodzian 2000], according to which concepts take precedence over words. In GTT terms are monosemic and concepts are monoreferential; a perfect correspondence is established between them: each concept corresponds to one and only one term. The term is thus equivalent to a symbol, to a "désignateur rigide" [Kripke 1995], since it works in much the same way as the system of proper names works in general language. Conceptual structures, on the other hand, are regarded as universal and static like their referents and any diachronic dimension is neutralised.

In outlining the theoretical and methodological differences between lexicology and terminology, Cabré (1999: 8) emphasized that terminology is only concerned with synchronic aspects:

This view, which today is considered the most systematic, coherent theoretical approach to terms, differs from lexicological theory in three ways: in the priority of the concept over the designation; in being exclusively concerned with the level of the terminological unit and not with the other levels of linguistic description; in excluding any diachronic approach of information.

Terminology began to open up to diachrony timidly after the "odyssée terminotique" [Slodzian 2000] of the 80s, when an impressive number of term banks, based on the synchronic traditional vision of the Viennese school, were produced and spread (EURODICAUTOM, NORMATERM, UNITERM, TERMIUM, IATE etc.). In 1988 a symposium entitled

“Terminologie diachronique” was organized by the *Centre de Terminologie* of Bruxelles, but ten years after Møller (1998) still regretted a “quantitative and / or qualitative deficit” in that branch of terminology for which he coined the neologism “terminochronie”. Over the last twenty years several studies have been conducted focussing on the evolution of concepts and terms in specialized languages (see inter al. van Campenhoudt 1997, Dury 2008, Zanola 2014, Kristiansen 2014), as well as on methods for compiling diachronic corpora [Dury 2004] and identifying the evolution of knowledge in corpora [Picton 2009].

As underlined by Dury and Picton (2009), this lack of interest for diachronic aspects in terminology was mainly due to technical, pragmatic, psychological, and most of all theoretical and historical obstacles. Diachronic terminology is, indeed, by definition a textual terminology [Bourigault and Slodzian 1999] [Pearson 1998], a terminology of “discourse”. Terms must be analysed in their real context of use and must not be seen as abstract and static labels associated with concepts. They should be rather regarded as linguistic signs, which enter into a dynamic relationship both with the concepts and with the other linguistic signs that constitute the terminological system. In a diachronic perspective, terminological work is fundamentally semasiological and descriptive. In order to grasp terminological variation over time, the starting point is the sense that a term acquires within a real linguistic context and the focus is on language as it is actually used (*Ist-Norm*) in real specialised (con-)texts and not on language as it should be used (*Soll-Norm*). To use Bourigault and Slodzian words (1999: 31), “le texte est le point de départ de la description lexicale à construire. On va du texte vers le terme”.

Texts are usually the unique source that the scholars have at their disposal in order to access disappeared states of languages and more completely understand the history of terms and their evolution over time. Texts play a crucial role to establish when and how words acquire (terminologization) and lose (determinologization) their terminological status, or how a term passes from one specialized field to another or from the intimate pages of a scholar to the entire scientific community.

Regarding this last aspect, in particular, texts constitute a precious means to record the typical terminological fluctuation that precedes or follows the formation of the so-called “normal science” state. As a matter of fact, in a phase of “normal science”, when the development of knowledge is adherent to what Kuhn (1970) calls “paradigm”, the system of terms is standard, explicitly defined and consensually used within a scientific community. In revolutionary phases, on the contrary, when a revision of existing scientific beliefs or practices is involved, terminology can be characterized by great instability. Texts become thus a conceptual and terminological space of creativity, where scholars think, reflect and give a linguistic form to their ideas. In these cases, ephemeral lexical formations are often attested only in the idiolect of the one who forged them without ever taking root within the scientific community.

An example of this phenomenon will be illustrated in Section 6 with Ferdinand de Saussure’s terminology.

3. Modelling an Interoperable Terminological Resource

When modelling a digital scholarly resource, the key point is to choose a model capable of maximizing its usefulness in the community of reference. According to the “FAIR guidelines”, scholarly data should be Findable, Accessible, Interoperable, and Reusable [Wilkinson et al. 2016]. First of all, for data to be findable, they must be referred to using a globally unique and persistent identifier and must also be indexed in a searchable resource. Furthermore, data must be accessible through a standardized protocol and, at the same time, has to be interoperable in virtue of its representation in a formal and shared knowledge representation language. Finally, to be reusable, data must be “described with a plurality of accurate and relevant attributes” (ibidem: 4).

In order to pursue the FAIR policies and to build fair diachronic terminological resources, our choice was to adopt the paradigms of Semantic Web technologies [Berners-Lee et al. 2001] and Linked (Open) Data [Bizer et al. 2011].

The intuition behind the Semantic Web is to shift from a Web of documents, viewed as separate data silos containing unstructured or semi-structured data, to a Web of data where open repositories of semantically enriched data can be accessed and interlinked. To make this vision concrete, the following technologies - often referred to as the Semantic Web layer cake or Semantic Web technology - are used: Resource Description Framework^[3] (RDF), RDF Schema

(RDFS), Web Ontology Language (OWL), Semantic Web Rule Language (SWRL) and SPARQL Protocol and RDF Query Language (SPARQL).

Semantic Web and Linked (Open) Data best practices require datasets to be published using RDF: data is described as triples in the form of Subject - Predicate - Object, where the Subject is a resource, the Object can be either a resource or a data value (such as a string of character or a number), and the Predicate (also called “property”) relates the Subject and the Object together. Basically, a dataset in the Linked Data world is a set of triples.

OWL is a family of languages, based on RDF/XML, and recommended by the World Wide Web Consortium since 2004 for representing and sharing ontologies^[4] on the Web. There are three variants of OWL, with different levels of expressiveness: OWL Lite, which supports a classification hierarchy and simple constraints; OWL DL, which provides the maximum expressiveness while retaining computational completeness and decidability, and OWL Full which allows for maximum expressiveness, resulting nonetheless undecidable.

In order to publish structured data on the Web in a way that they can be easily linked to each other, a number of rules was set up by Berners-Lee, known as Linked (Open) Data principles. According to them, each entity of a dataset must be uniquely identified by a special “address” called URI (Uniform Resource Identifier), each URI must be available on the Web via the HTTP protocol^[5], each look up of a URI must be answered with useful information and using standards, and finally, whenever possible, data should be linked to other data.

Over the last years, the Linked (Open) Data approach has attracted the attention of a linguistic community becoming more and more sensitive to the question of interoperability between resources. An increasing number of openly available and interconnected linguistic datasets have been published as part of the Web of data, resulting in the so-called Linked Open Data Cloud (LOD). This trend, nicely visualised in <https://lod-cloud.net>, is particularly boosted by the Ontology Lexicon (Ontolex) community group, whose activity resulted in the creation of *lemon*^[6] (Lexicon Model for Ontologies), a meta-model allowing lexical resources to be published in RDF, linked and shared according to Semantic Web principles. At present *lemon* constitutes a standard *de facto*, as it is grounded on W3C and ISO standards, such as LMF (Lexical Markup Framework) [Francopoulo et al. 2006], Lexinfo [Cimiano et al. 2011] – this latter aligned with ISOCat - and LIR (Linguistic Information Repository) [Montiel-Ponsoda 2008].

Lemon has been successfully adopted in many different contexts for different purposes, such as, for example, to publish several existing lexical resources as linked data, including DBnary [Sérasset 2015], the Parole-Simple-Clips Italian lexicon [Del Gratta et al. 2015], WordNet [McCrae et al. 2014], UBY [Eckle-Kohler 2015] and BabelNet [Ehrmann et al. 2014], to represent linguistic resources for sentiment analysis within the EuroSentiment project [Sánchez-Rada et al. 2014], to model linguistic annotations in FrameBase, a linked open knowledge base which integrates and interconnects heterogeneous sources of structured knowledge [Rouces et al. 2017].

The *lemon* core module, named *ontoLex*^[7], is composed of three fundamental elements: i) the *Lexical Entry* class, in turn specified in *Word*, *Multiword Expression*, and *Affix*; ii) the *Lexical Form* class where the morphological variants of a lexical entry are described and associated with a written representation (*writtenRep*); iii) the *Lexical Sense* class, which constitutes the reification of the relationship between a lexical entry and an ontology entity defined in a given “external” ontology. In *lemon*, in fact, concepts, conceived as extra-linguistic entities, receive a formal and explicit description in a separate ontology, according to the principle known as “semantics by reference” [Buitelaar 2010].

The *lemon* model has been designed to be modular and extensible: in other words, new modules can be created, if required, to represent a peculiar aspect of a lexical or terminological resource. To manage semantic shifts, for example, a diachronic extension, called *lemonDIA*, was developed [Khan et al. 2014].

4. Models and Tools for the Representation of Diachrony

Knowledge representation languages underlying Semantic Web technologies are based on binary relations. They typically connect two instances without any temporal information, thus making the representation of time a difficult matter to deal with. In order to tackle this issue, different approaches have been proposed in the literature (see inter al.

Flouris 2008, Makri 2011): Temporal Description logics [Artale and Franconi 2000] [Lutz et al. 2008], Versioning [Klein and Fensel 2001], Reification [Kanmani 2016] and 4D-fluents [Welty et al. 2006]. In the following, a brief description of each approach is provided.

Temporal Description Logics (TDL). Description logics (DLs) [Baader et al. 2003] [Baader and Sattler 2001] [Calvanese et al. 2001] are a family of knowledge representation formalisms underlying the Ontology Web Language (OWL-DL). They are a decidable subset of first-order logic. TDLs extend DL with standard temporal logic operators such as “always in the future”, and “until”. Many TDLs have been proposed in the literature [Artale and Franconi 2000] [Lutz et al. 2008] with the most expressive of them being undecidable. Extending DL entails to extend OWL, which is based on DL. Since OWL is a W3C recommendation (2004) for publishing and sharing ontologies in a “Semantic Web vision”, adopting TDL would entail to abandon the Semantic Web framework. On the contrary, the other approaches described in the following can be implemented using OWL directly.

Versioning. Another strategy commonly adopted to handle and keep track of ontology modifications in time is versioning: whenever a change takes place, a new version of the resource is created (see Figure 1). Several works are grounded on this approach: Grandi (2009) presents a multi-temporal RDF database model employing triple timestamping with temporal elements; Zekri et al. (2017) present an infrastructure and a suite of tools to support the creation and validation of OWL ontologies with time-varying instances; in Tappolet and Bernstein (2009), the usage of Named Graphs is proposed.

The drawback of versioning approach is mainly twofold. Firstly, the fact that a new version of the whole resource needs to be created every time a change occurs leads to redundant information. Secondly, the more numerous the versions of the resource are, the more complex and time-consuming the performance of exhaustive searches becomes.

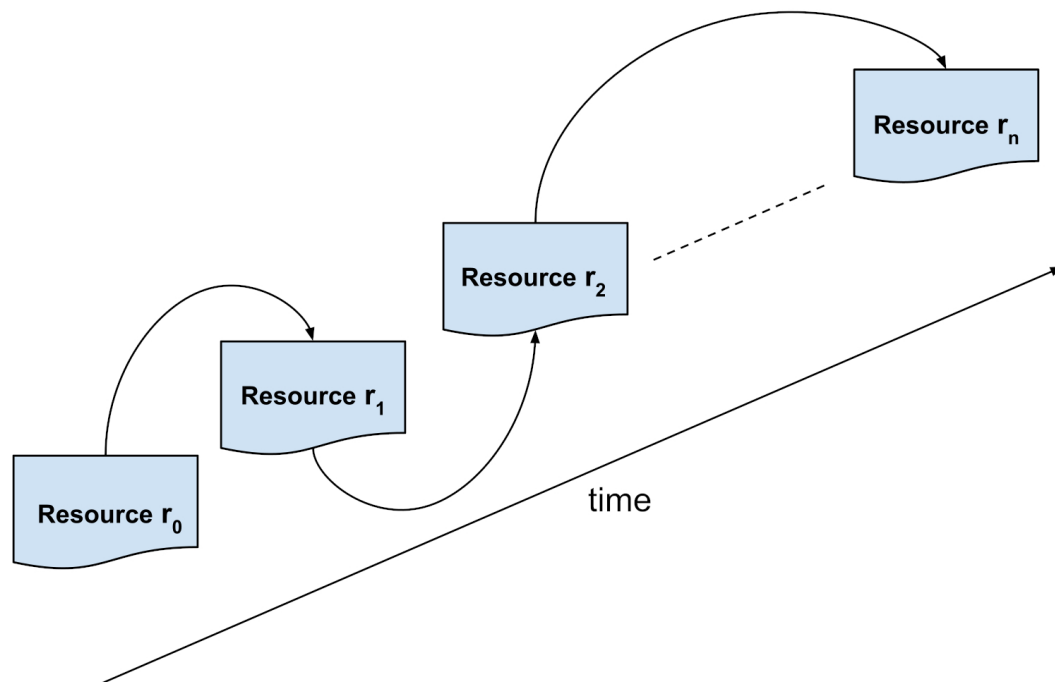


Figure 1. *Versioning* - when a change takes place, a new version of the resource is created.

Reification. Reification is a technique for representing relationships with arity greater than two.^[8] Reifying a relationship means representing it as a class of entities. The N-ary Relations pattern, conceived to model a relationship as a resource, is based on reification.

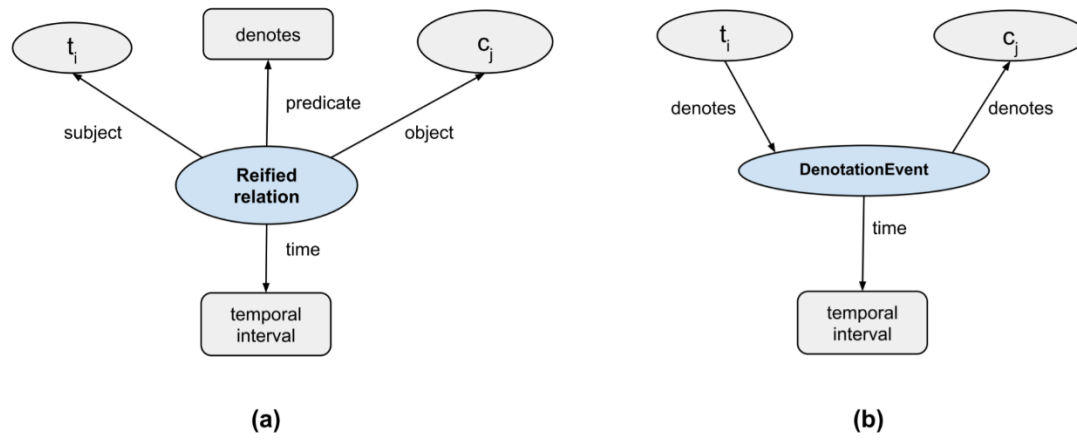


Figure 2. Example related to the fact that the sense of the term t_i denotes the concept c_j : (a) example of reification - (b) example of N-ary relation pattern.

Figure 2(a) illustrates how the reification process works: the relation *denotes* is transformed into a class, here named *Reified relation*. The subject and the object of *denotes*, as well as the relation itself, become the attributes of the new class. Let us suppose that the relation *denotes* links a term t_i (in particular its specific sense) to a concept c_j at a specific time T_0 . As shown in Figure 2(b), the relationship *denotes*(t_i , c_j , T_0) can be reified as the new class *DenotationEvent* with three binary relations: *denotes*(t_i , d_e), *denotes*(d_e , c_j), and *time*(d_e , T_0). The first two relations link t_i to c_j by means of d_e , which is an instance of the class *DenotationEvent*. The third relation assigns a temporal extent T_0 to d_e , representing the temporal validity of the original property *denotes*(t_i , c_j).

The main disadvantages of this technique are that: i) the complexity of the ontology increases due to the proliferation of entities, both classes and properties (as a matter of fact, whenever a temporal relation has to be represented, a new object is created); ii) the OWL reasoning capabilities become limited [Welty et al. 2006] and can be maintained only by means of time-consuming strategies, since the reified property is transformed into a class and the characteristics which can be added to a property in OWL (such as functionality, inverse functionality, transitivity or symmetry) are no longer associated directly with the relation itself.

4D-fluents (four-dimensionalist). The basic idea underlying this approach is that each entity has a temporal part, described by a set of time slices. Each slice represents a temporal period, in which all the features of an entity remain constant. Changes affect the properties of the temporal part keeping the entity unchanged. In order to add the time dimension to a resource, the *TimeSlice* and *TimeInterval* classes with the *timeSliceOf* and *timeInterval* properties need to be introduced, as Figure 3 depicts. Returning to the previous example, in the 4D-fluents approach, the property *denotes* would connect instances of the *TimeSlice* class and would hold between slices of entities. The time slices of an entity have a specific lifetime, that is the time interval of the relation they participate in.

As in the N-ary model so also in the 4-D fluents approach the domain and the range of the property *denotes* change, both represented by the *TimeSlice* class. So, this model as well suffers from some limitations in OWL reasoning capabilities. With respect to the N-ary model, the 4-D fluents approach is affected by a wider proliferation of entities (compare the blue objects in Figure 2).

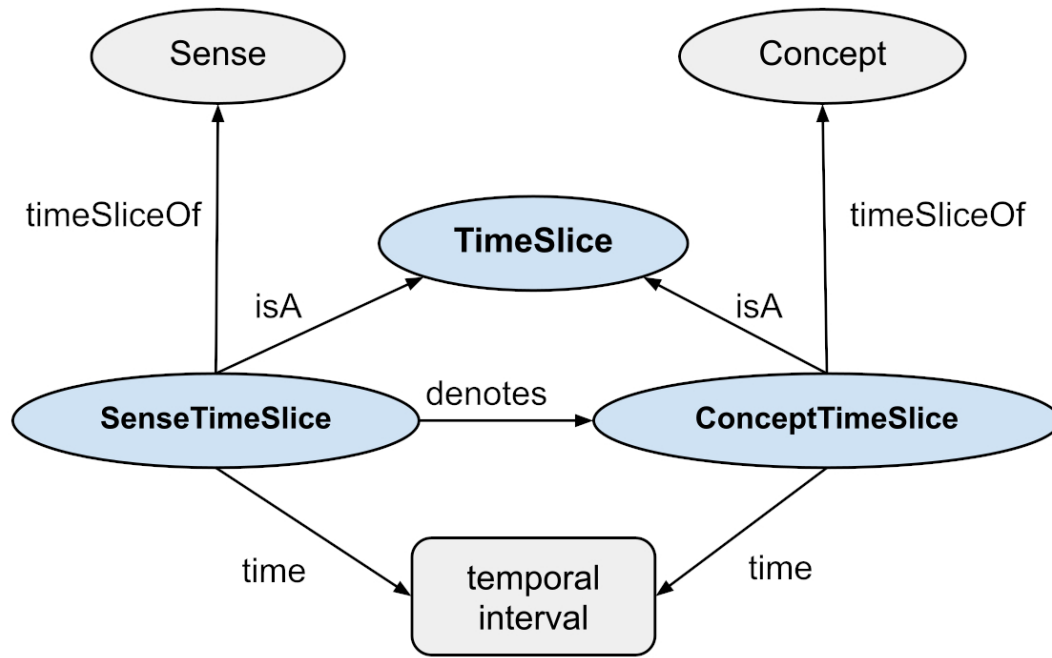


Figure 3. Example of 4D-fluents representation.

So far, the main mechanisms for representing time evolving information were illustrated. In the next two subsections, we present a vocabulary to represent time-related facts, i.e. OWL-Time, and a set of rules, i.e. SWRL, which operate on temporal relations to infer new temporal knowledge and to detect inconsistent assertions.

4.1 OWL-Time and Allen's Relations

OWL-Time^[9] is an ontology of temporal concepts, conceived to describe the temporal properties of resources. It provides a vocabulary for expressing facts about topological relations among instants and intervals, together with information about durations and date-time. Figure 4(a) illustrates a diagram of OWL-Time ontology. The properties *hasTemporalDuration*, *hasBeginning*, *hasEnd*, and *hasTime* provide a standard way to attach time information to ontological entities.

Let us consider, for example, that the term t_i denotes the concept c_j from 1898 to 1903. In this case all the needed temporal facts can be specified using OWL-Time vocabulary.

However, if we wanted to answer questions like "What are the intervals overlapping the interval 1898-1903?", or "what are the relations holding in that interval?", OWL-Time vocabulary would not be enough. James F. Allen (1983) provided then a calculus for temporal reasoning, by introducing thirteen base relations capturing the possible relations between two intervals (see Figure 4(b)). They allow the computation of any relative position or sequence. In the context of the Semantic Web, these rules can be represented by means of the Semantic Web Rule Language (SWRL), described in the following section. The use of this language makes it possible to assert new instances of relationships between the ontology entities.

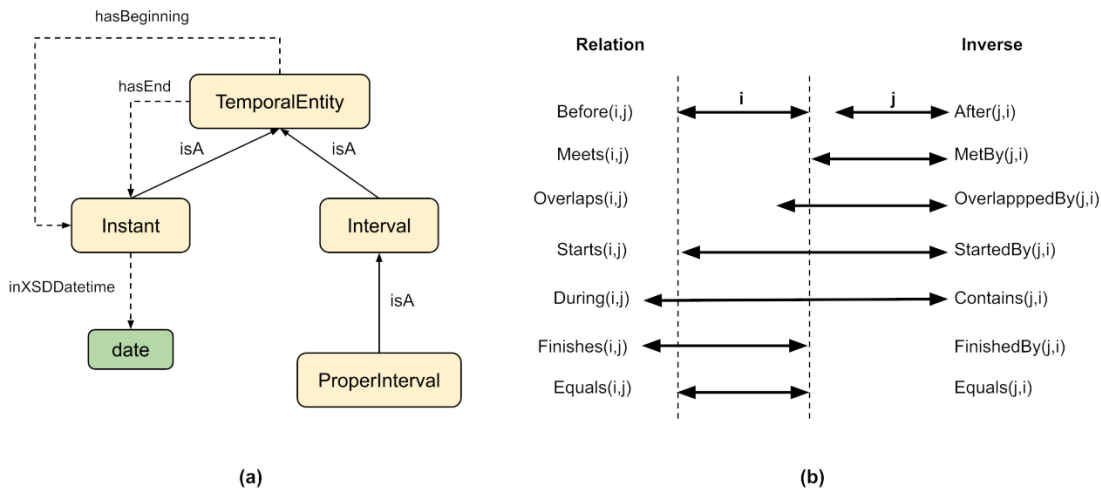


Figure 4. (a) OWL-Time - the core model of temporal entities. (b) Allen's rules between time periods.

4.2 Semantic Web Rule Language

Semantic Web Rule Language^[10] (SWRL) is a language based on a subset of Datalog with both unary and binary predicates, representing the rule layer within the Semantic Web stack. The form of these rules is an implication between an antecedent (body) and a consequent (head), each consisting of zero or a conjunction of more atoms.^[11] Whenever the conditions specified in the antecedent hold, then the conditions specified in the consequent must hold as well. Atoms can be of the form $C(x)$, $P(x,y)$, $\text{sameAs}(x,y)$, $\text{differentFrom}(x,y)$, or $\text{builtIn}(r,x,...)$ where C is an OWL description or data range, P is an OWL property, r is a built-in relation, x and y are either variables or OWL individuals or OWL data values, as appropriate. For example, the following SWRL rule infers if an instant i_1 is before an instant i_2 , by verifying a set of premises:

$\text{Instant}(?i_1) \wedge \text{Instant}(?i_2) \wedge \text{inXSDDatetime}(?i_1, ?y) \wedge \text{inXSDDatetime}(?i_2, ?w) \wedge$
 $\text{swrl:lessThan}(?y, ?w)$
 $\rightarrow \text{before}(?i_2, ?i_1)$

where the built-in $\text{swrl:lessThan}(?y, ?w)$ is satisfied when the date represented by the variable $?y$ is earlier than that represented by $?w$.^[12] In this case, when temporal information is provided as a date, the qualitative relations are specified using SWRL rules that apply on the quantitative representation. The rules can be applied to qualitative temporal information as well. In the following, a rule about the relation *during* is provided:

$\text{ProperInterval}(?a) \wedge \text{ProperInterval}(?x) \wedge \text{before}(?b, ?y) \wedge \text{before}(?z, ?c) \wedge$
 $\text{hasBeginning}(?a, ?b) \wedge \text{hasBeginning}(?x, ?y) \wedge \text{hasEnd}(?a, ?c) \wedge \text{hasEnd}(?x, ?z)$
 $\rightarrow \text{During}(?x, ?a)$

5. Adding Time to Terminologies

Given the different approaches to modelling time described in the previous section, we chose to adopt the N-ary model.

The reason for this choice is twofold. First of all, the N-ary relations approach is supported by the Ontology Engineering and Patterns Task Force of the Semantic Web Best Practices and Deployment Working Group.^[13] Secondly, with respect to the perdurantist approach, the N-ary model requires the introduction of a smaller number of objects and thus "outperforms the 4D-fluents representation in terms of required assertions (axioms) and consequently in reasoning time" [Batsakis et al. 2017, 14].

In order to formalize terminology evolution following the N-ary approach, we took inspiration from the model introduced

by Preventis et al. (2014)^[14] to handle temporal ontologies in OWL. In particular, we introduced the following entities:

- the class *Event*, which represents the reification of a property holding in a certain time. To model the class *Event* we used the relative type *dcmitype:Event* taken from the Dublin Core Metadata Initiative (DCMI) and defined as a “non-persistent, time-based occurrence”^[15];
- the object property *during*, having the class *dcmitype:Event* as domain and the class *time:ProperInterval* of the OWL-Time ontology as range. The latter is defined as “a temporal entity with non-zero extent or duration, i.e. for which the value of the beginning and end are different”.^[16] *during* has been modelled as a sub-property of *dcterms:date* (described as a “point or period of time associated with an event in the lifecycle of the resource”) of the DCMI vocabulary^[17];
- the object property *temporalProperty*, which relates individuals of the class *dcmitype:Event* respectively to the source and to the target individuals of the original property which has been reified. As we will see in detail shortly, the object properties that are converted into temporal ones are formalised as sub-properties of *temporalProperty*.

Following the N-ary approach outlined above, the terminological layer of our model was formalised by extending the *lemon* model. The model does not contain a complete collection of linguistic categories; consequently, it builds on Lexinfo ontology^[18] in order to provide a vocabulary to describe the properties of linguistic objects. More specifically, the four lexico-semantic relationships defined in Lexinfo, namely synonymy, antonymy, hypernymy, hyponymy^[19] were “temporalised” (*temporalSynonym*, *temporalAntonym*, *temporalHypernym*, and *temporalHyponym*) and thus converted into sub-properties of *temporalProperty* to indicate their involvement in a temporal event. The same was done for the properties *ontolex:reference*, which links every individual of the *ontolex:LexicalSense* class to a concept defined in an ontology, and *ontolex:sense*, linking an individual of the class *ontolex:LexicalEntry* to a specific sense.

In addition, each of these relations was reified and converted into subclasses of the class *dcmitype:Event*: *senseEvent* representing the reification of sense relation, *referenceEvent* representing the reification of the relation linking a sense and an ontology concept, *equivalentEvent*, *incompatibleEvent*, *broaderEvent*, *narrowerEvent* representing respectively the reification of synonymy, antonymy, hypernymy, hyponymy relations holding between two senses.

The class *dcmitype:Event* as well as its subclasses mentioned above are related to the class *time:ProperInterval* through the object property *during*.

So, for example, in order to represent the temporal synonymy (Figure 5), the class *equivalentEvent* is related to the class *time:ProperInterval* through the relation *temporalSynonym*, specifying the time period in which the relation of synonymy holds between two senses.

As we can see, the introduction of an intermediate object, namely the individual of the class *equivalentEvent*, splits the static relation linking a subject and an object into two parts. Instead of a single triple linking two senses (s_1 sense s_2), two triples are created, in which the individual of the class *equivalentEvent* appears as subject in one case and as object in the other.

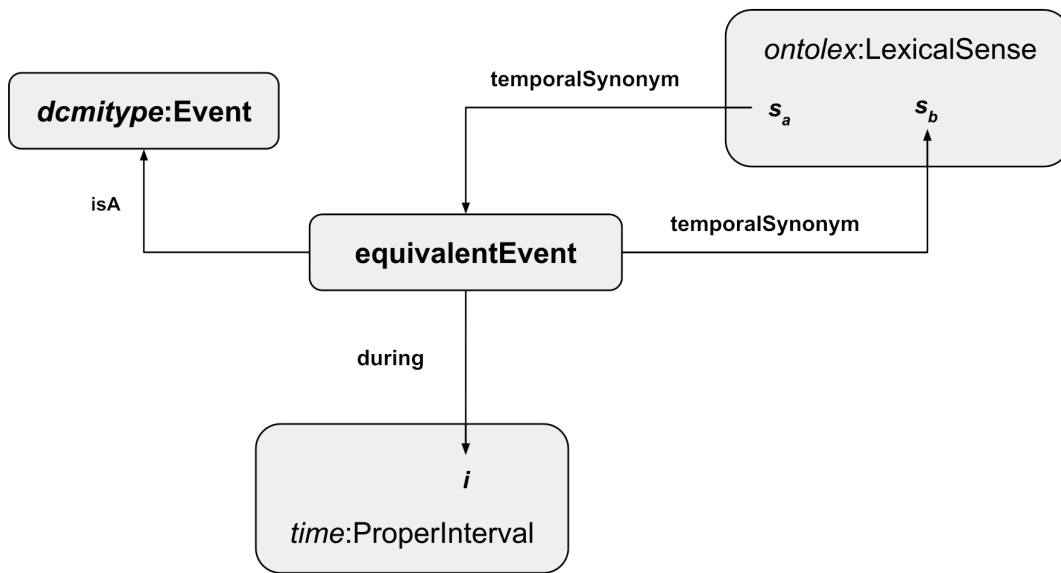


Figure 5. Synonymy is represented as a temporal event that occurs during an interval: sense s_a of a certain lexical entry is asserted to be synonym of sense s_b of another lexical entry in a specific time interval i .

However, this is not enough. As we underlined above, to define in which period a term is used with a certain sense and to denote a specific concept, scholars need to resort to textual sources.

52

As a result, our model is based on the key concept of attestation; more specifically, the senses of a term are linked to the text where they are attested, and for each text the writing time is defined by means of specific relations. To represent texts and attestations, as the main entities constituting the textual layer of our model, we introduced the following entities:

53

- the class *Text*, which represents every work produced in a written form. As many methodological and theoretical works have underlined over the years, the notion of text is highly complex and culturally rich and, thus, difficult to define and formalize. A detailed discussion is beyond the scope of this article. The term “text” is to be understood here in a broad sense, including entities that are very different from each other, such as books, published articles, handwritten notes, drafts of unfinished works etc. It can be considered equivalent to the type “Text” as defined in the Dublin Core Metadata Initiative Type Vocabulary (*dcmitype:Text*), namely a “resource whose content is primarily words for reading”.^[20] If scholars need to use vocabularies providing more subtle distinctions, many models for representing documents in Linked Data have been already proposed, offering an opportunity for high quality bibliographic data to be exposed to the Semantic Web, such as FRBR (Functional Requirements for Bibliographic Records) [Carlyle 2011] or Bibframe (Bibliographic Framework) [Kroeger 2013]. Importantly enough, if a different vocabulary is adopted, the validity of our model is not compromised;
- the object property *hasAttestation*, which relates a sense, instance of the class *ontolex:LexicalSense*, to an instance of the class *dcmitype:Text*; a few works about the modeling of the attestation of a term in a document have been already done [Khan et al. 2017] [Bellandi et al. 2017]; for the sake of reuse of existing vocabularies we adopted the *hasAttestation* property defined by the LAWD (Linked Ancient World Data) ontology^[21];
- the object property *hasWritingTime*, defined as a subproperty of *dcterms:date*, which relates an individual of the class *dcmitype:Text* to an individual belonging to the class *time:ProperInterval* of the OWL-Time ontology.

It is up to the scholar to define the boundaries of the writing interval, which can be qualitative or quantitative. In fact, only in rare cases he/she can define with certainty the extent of the period in which an author has worked on drafting a text,

with the aid of internal elements or external sources (for example, the author explicitly declares that he/she has written that work in a certain period, etc.).

In most cases, on the contrary, temporal information on the writing process is vague and ambiguous and partially bounded intervals (“before X”, “after Y”, etc.) need to be defined. Instead of specifying the exact duration or the starting and ending points in time, scholars describe intervals with respect to their mutual relations, by establishing the *terminus post quem* or the *terminus ante quem*, namely the limits of the possible range of dates in which a work has been written. [22]

As for published texts, instead of the *hasWritingTime* property, the *dcterms:issued*^[23] can be used, with the year of publication constituting the end of the drafting period.

An overview of the model is illustrated in Figure 6.

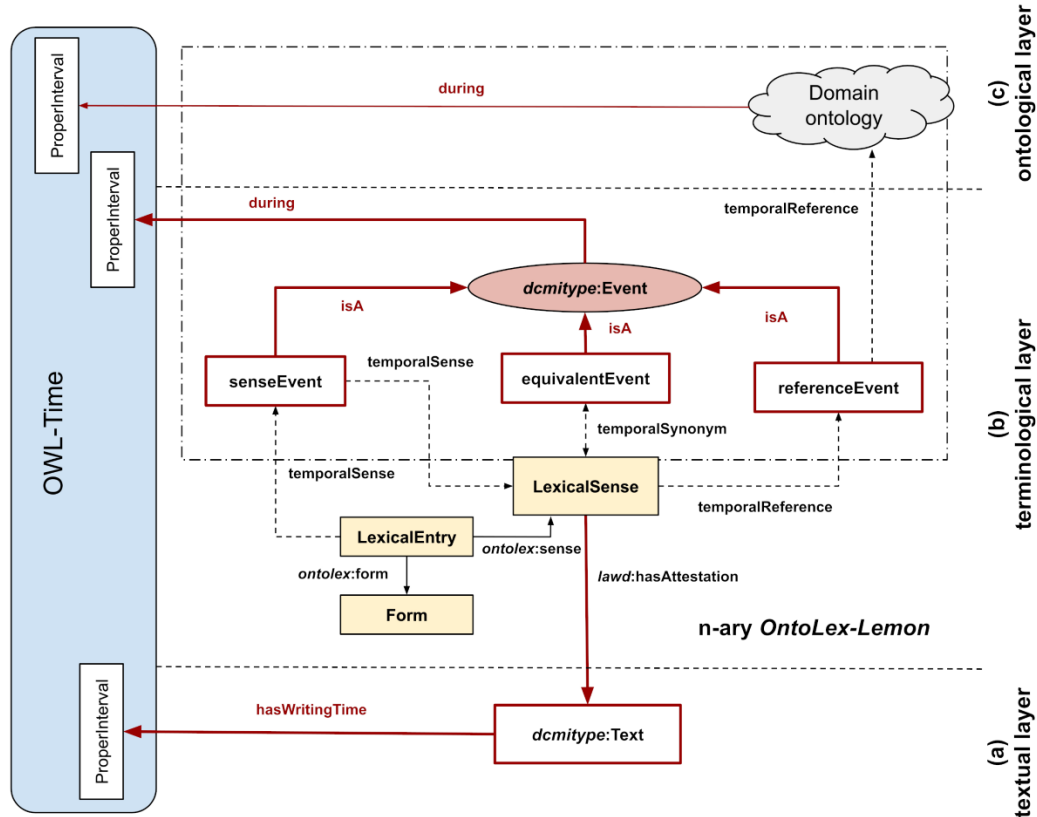


Figure 6. Overview of the proposed model. (a) the textual layer is constituted by texts and attestation. (b) the terminological layer relies on a N-ary version of *lemon*, where each reified class is an event having a temporal extent. (c) the ontological layer concerns the formal definitions of the concepts denoted by the terms and it is described with OWL.

To take maximum advantage of the adoption of OWL as the representation language of the resource, we have also conceived a set of SWRL rules to automatize the attribution of some temporal information.

The following rule states that, if a sense is attested in a text written in a certain time interval, then that sense “exists” during the same interval.

```
Text(?t) ^ time:ProperInterval(?i) ^ hasWritingTime(?t, ?i) ^ ontolex:LexicalEntry(?l) ^ ontolex:LexicalSense(?s)
^ isAttestedIn(?s, ?t) ^ ontolex:sense(?l, ?s) ^
swrlx:createOWLThing(?se, ?s)
→ senseEvent(?se) ^ temporalSense(?l, ?se) ^ temporalSense(?se, ?s) ^ during(?se, ?i)
```

More specifically, given a text t written in a time interval i (defined as an instance of *time:ProperInterval*), and given a lexical entry l with a lexical sense s (these two related to each other through the synchronic *ontolex:sense* relation) attested in text t , then l has sense s (in a diachronic perspective) during the interval i . The premise of the rule is composed of the following steps: i) an anonymous individual se is created *ex novo* in the antecedent part of the rule using the *createOWLThing* built-in relation^[24]; ii) se is defined, in the consequent part, as an event representing the reified relation (in particular by instantiating it as a member of the class *senseEvent*); iii) the property *temporalSense* is used to relate the lexical entry l to se and, in turn, to relate se to sense s ; iv) finally, the event se is defined as happening in the same time interval i during which the text has been written.

60

If the sense of a lexical entry is attested in more than one text (each one written in a specific temporal interval), the rule will be fired for each text, thus relating the lexical entry to that sense in a more complex period constituted of multiple time intervals.

61

Similarly, the temporal validity of the *ontolex:reference* relation, linking each *ontolex:LexicalSense* to one and only one concept of an ontology, can be inferred from the writing time of the text as well, using the following SWRL rule:

62

```
Text(?t) ^ time:ProperInterval(?i) ^ hasWritingTime(?t, ?int) ^
ontolex:LexicalSense(?s) ^ isAttestedIn(?s, ?t) ^ Concept(?c) ^ ontolex:reference(?s,
?c) ^ swrlx:createOWLThing(?re, ?c)
→ referenceEvent(?re) ^ temporalReference(?s, ?re) ^ temporalReference(?re, ?c) ^
during(?re, ?i)
```

63

If a sense s is linked to a concept c through the (synchronic) *ontolex:reference* property, and if s is attested in a specific text t written in the time interval i , then, in a diachronic perspective, the relation *reference* holds in the same time interval i as well. In other words, the fact that a word is used with a particular sense in a specific text is equivalent to saying that in that period that term was used to express that specific concept.

64

Figures 7 and 8 show respectively how the temporal versions of *ontolex:sense* and *ontolex:reference* can be automatically inferred by a reasoner and represented in our N-ary model starting from the writing time of texts.

65

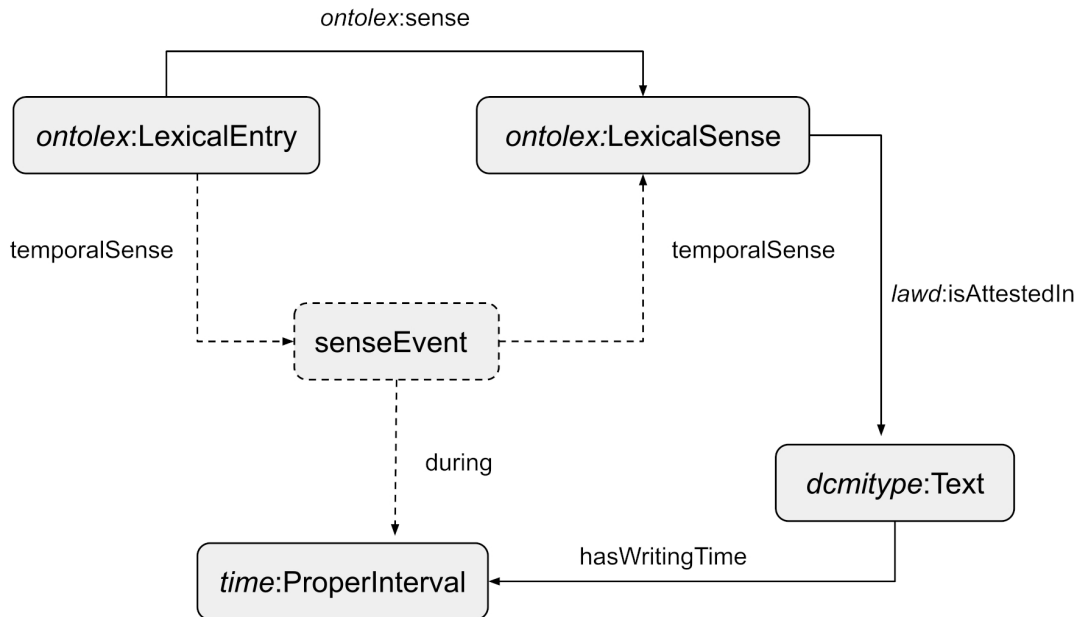


Figure 7. Classes and properties appearing in solid lines constitute the synchronic part of the resource. Elements in dashed lines, representing the temporal version of the *ontolex:sense* relation, can be automatically implied by the firing of the SWRL rule.

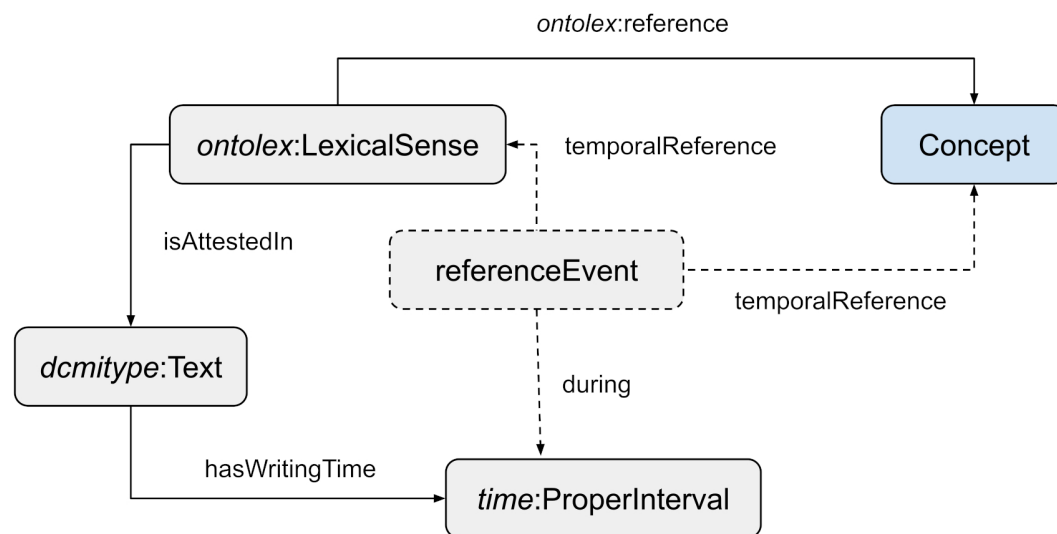


Figure 8. Temporalization via reification of the *ontolex:reference* relation.

We are also working at defining other rules taking into account the relations of synonymy, antonymy, hypernymy, and hyponymy, with the aim of automatizing the creation of their diachronic counterparts, with the addition of Allen's relations described in Section 4.1.

66

6. The Construction of a Diachronic Resource: the Saussure Case Study

Ferdinand de Saussure (Geneva, 1857-1913), the father of general linguistics, never published his theories on general linguistics and semiotics in an organic work. He strove to structure the flow of his thoughts and ideas in handwritten notes, which are often characterized by a synthetic, fragmentary and hermetic nature. These manuscripts display an obsessive and constant search for clear, effective and unambiguous terms, against the "ineptie de la terminologie courante" (Letter to Meillet written in 1984, see Benveniste 1964). He changed the meaning to some terms over time, assigned additional specific meanings to already existing terms, used some expressions ephemerally, and even forged new words. This terminological fluctuation reflects the difficulties underlying every process of theoretical creation and sheds light on the evolutionary dynamics through which concepts take shape and cut through language to finally find their linguistic expression. This process of *rumination* [Fenoglio 2012] is typical when dealing with theoretical terminologies as in the case of Saussure, who tries to revolutionize the current linguistic vision and to lay the foundations of linguistics as a science.

67

Saussure's terminology has been the subject of numerous studies [Godel 1957] [Engler 1968] [Cosenza 2016]. In 2011 the first electronic thesaurus of Saussure's terminology, named *Simple_FdS* [Ruimy et al. 2013], was built as part of the project "Per una edizione digitale dei manoscritti di F. de Saussure", coordinated by then President of the Cercle Ferdinand de Saussure Daniele Gambarara.

68

The architecture of *Simple_FdS* was based on the lexical model SIMPLE [Lenci et al. 2000] and was therefore inspired by the Generative Lexicon theory elaborated by Pustejovsky (1995). The Generative Lexicon theory, mainly applied to general language, has proved to be very effective in the description of specialized language and terminological lexicons as well (for example Aráuz et al. 2012, Sánchez Ibáñez and García Palacios 2014), as it makes it possible to enrich the range of conceptual relations which is traditionally based on generic-specific and part-whole relations.

69

Initially, the electronic lexicon was conceived as static and no diachronic aspects were taken into account. Recently, as we shall illustrate in Section 6.1, the model here proposed has been adopted to convert this terminological resource from a static repository into a dynamic one.

70

As underlined in Section 2, in diachronic terminology the textual dimension plays a crucial role. Consequently, the first

71

step was to build a good corpus, able to represent the diachronic evolution we want to study. The texts included in Saussure's corpus can be considered as highly representative of diachronic evolution of Saussure's terminology, embracing all the thirty-seven years in which the intellectual activity of the Genevan linguist unfolded. The period covered by the corpus ranges from 1874 - when Saussure wrote his first *Essai pour réduire les mots du grec, du latin & de l'allemand à un petit nombre de racines* - until 1911, the year in which he took the third course of general linguistics, the last academic teaching held before his death, which occurred in 1913 after a long illness.

Let us see in detail an example of the formalisation of the terminological changes occurring in Saussure's writings.

72

6.1 Temporalizing Saussure's Terminology: the Case of *Signe*

As is well known, Saussure's theories have been particularly influential in semiotic studies. He emphasized first the dyadic and complex nature of linguistic sign, which consists of two heterogeneous entities: a signifier (*signifiant*), namely the hearer's psychological impression of a sound, and a signified (*signifié*), i.e. a concept, an abstract idea. This terminology - still in use today - was introduced by the Genevan linguist at the end of his life, after a long theoretical reflection emerging in many of his handwritten pages.

73

Here, as an example, the concept of SIGNIFIER was chosen and the formal representation of the different terms Saussure used over time to designate this concept is illustrated, in accordance with the N-ary model we propose here. In the following, the relationships and the classes taken from existing standard vocabularies will be indicated with prefixes of the respective vocabularies.

74

The concept of SIGNIFIER is denoted by many different terms in Saussurean manuscripts, such as *signe*, *image acoustique*, *image verbale*, *image vocale*, *image auditive*, *son*, *forme*. More specifically, the term *signe* is attested intermittently from 1891 to 1911 in many works, such as *De la double essence du langage*, *Status et Motus*, *Notes Item*, *Troisième Cours de linguistique générale* etc. To formally represent this, the instance *signe_sense*, belonging to the *ontolex:LexicalSense* class, was created and linked - by means of the relation *isAttestedIn* - to the instances of the class *dcmltype:Text* representing all the works where the term occurs. The writing period of each text was formally defined. Here the case of *Double Essence du langage* is shown. The instance *Double_Essence_Writing_Time* of the class *time:ProperInterval* (subclass of *time:Interval*) was created and linked to the instance *Double_Essence* through the relation *hasWritingTime*. The writing interval of this work, which is supposed to range from 1891 to 1892,^[25] was formally defined by linking the instance *Double_Essence_Writing_Time* through the relation *time:hasBeginning* to an instance of the class *time:Instant*, defined by the Data Property *XSDDateTime* with value "1891-01-01T00:00:00" as well as through the relation *time:hasEnd* to an instance of the class *time:Instant*, defined by the Data Property *XSDDateTime* with value "1892-12-31T00:00:00". The same formalisation was performed for each work where this sense of *signe* was attested.

75

As described in Section 5, on the basis of the N-ary approach here adopted, *ontolex* relations were reified and the following classes were introduced as subclass of *dcmltype:Event*: *broaderEvent* (reification of the hypernym relation); *narrowerEvent* (reification of the hyponym relation); *incompatibleEvent* (reification of the antonymy relation); *equivalentEvent* (reification of the synonymy relation), *referenceEvent* (reification of the inter-level relation *reference*) and finally *senseEvent* (reification of the relation linking a lexical entry to its senses).

76

By firing the two rules illustrated in Section 5, the classes *referenceEvent* and *senseEvent* were automatically populated. More specifically, in accordance with the first rule, the instance *signe_sense_ts* was created and linked by means of the relation *temporalSense* both to *signe_sense* (individual of the *ontolex:LexicalSense* class) and to *signe_entry* (individual of the *ontolex:LexicalEntry* class). Similarly, by firing the second rule, the individual *signe_sense_ref_signifiant_concept* was created and linked by the relation *temporalReference* both to *signe_sense* and to an ontology concept, described in the *Simple_FdS* ontology.

77

As underlined above, a terminological fluctuation can be observed in Saussure's work as regard to the designation of the concept SIGNIFIER.

78

An attempt to completely renew semiotic terminology is attested in the so-called *Notes Items*. In these handwritten pages, defined by Stetter (1979) as a “fragment sémiologique central”, Saussure’s reflection upon linguistic sign achieved extraordinary theoretical heights. Many heterodox neologisms were introduced, such as *sème*, *sème linguistique*, *apôsème*, *kenôme*, *sôme*, *contre-sôme* etc. Here the “material” part of linguistic sign, namely the “signifier”, is denoted in certain passages by the term *apôsème*. This terminology is abandoned later and the term *signe* continues to be used to indicate at the same time the “material” part of the linguistic sign and the whole, namely the combination of SIGNIFIED (concept) and SIGNIFIER (acoustic image). In order to avoid the confusion that the polysemy of the term *signe* generated in his students, Saussure introduced *in extremis* the term *signifiant*, during his last course on general linguistics in 1911. He proposed thus to retain the word *signe* to designate the linguistic entity as a whole, and to replace the dyad *signe* and *concept* respectively by *signifiant* and *signifié*.

79

To formalise this terminological fluctuation, the instance *signe_sense_syn_apôsème_sense* of the *equivalentEvent* class was created and the static triple (*signe_sense* *lexinfo:synonymy* *apôsème_sense*) was split into the following two triples in accordance with the N-ary model: (*signe* *temporalSynonym* *signe_sense_syn_apôsème_sense*) and (*signe_sense_syn_apôsème_sense* *temporalSynonym* *apôsème*). The instance *signe_sense_syn_apôsème_sense* was then linked to the interval *Notes_Item_Writing_Interval* by means of the relation *during*.

80

In a similar way, the instance *signe_sense_syn_signifiant_sense* of the *equivalentEvent* class was created and linked via *temporalSynonym* to the individual *signifiant*. The instance *signe* was in turn linked - by means of the same relation - to the instance *signe_sense_syn_signifiant_sense*, belonging to the reified class *equivalentEvent*. Finally, the instance was linked to the interval *troisième_Cours_Interval* (with beginning 1910 and end 1911) by means of the relation *during*.

81

6.2 Querying Saussure's Diachronic Terminology

Such a formalization, even if complex and redundant, makes it possible to reconstruct on a timeline the evolution of the lexical-semantic relationships between senses as well as the relations holding between senses and concepts, and then to answer complex and sophisticated queries. Here are just a few of the many queries which can be performed:

82

- Which are the terms used over time to denote a concept *c*?
- Which are the terms attested in a specific time interval *i* and used to denote a concept *c*?
- Which are the terms attested in a specific time interval *i* which are synonym (antonym or hypernym or hyponym) of the term *t*?
- Which are the terms synonyms of the term *t* in a specific text *T*?
- Which are the concepts denoted by the term *t* over time?
- Which are the concepts denoted by the term *t* in a specific time interval *i*?

Returning to our Saussure’s case study, let us suppose, for example, that scholars would like to know which terms have been used by the Genevan linguist over time as synonyms of the term *signe* (in the sense of *signe linguistique*, namely “l’union de l’idée avec ce produit phonatoire”).

83

The query results are represented here in the form of a graph (Figure 9), as they appear in LexO, a collaborative web editor of lexical and termino-ontological resources, developed by the Institute of Computational Linguistics of the Italian National Research Council (ILC-CNR) and already used within several research projects, such as DiTMAO [Bellandi et al. 2018] and Totus Mundus [Piccini 2018].^[26]

84

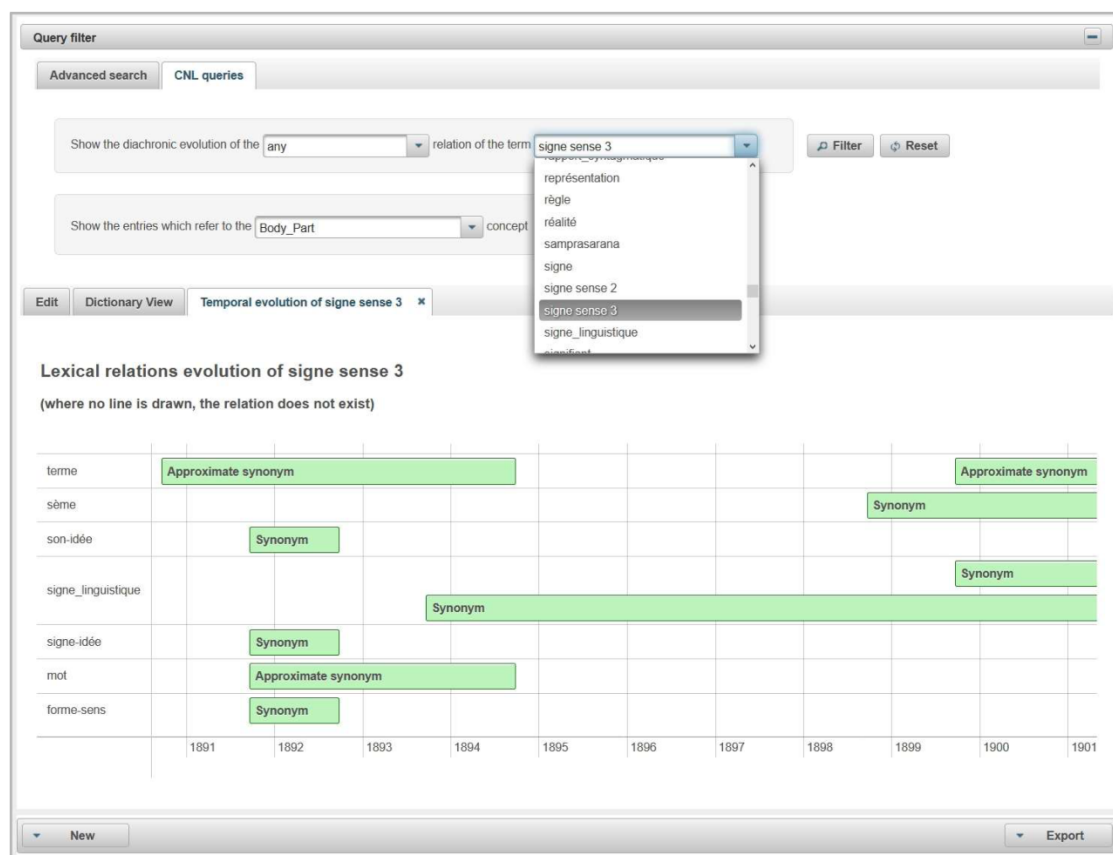


Figure 9. The synonyms of the term *signe* used by Saussure between 1890 and 1911

As clearly appears in the graph, the formalization previously described makes it emerge immediately Saussure's unceasing and obsessive work in search for the term that best represented the concept.

In an initial period, spanning approximately from 1891 and 1892, Saussure introduces three interesting neologisms: *signe-idée*, *son-idée* and *forme-sens*. These three nominal compounds well reflect the fact that the linguistic sign is composed of two parts that are closely connected as the two sides of a piece of paper. In effect, the bipartite structure of the terms mirrors the dyadic nature of the linguistic sign. Each element of the compound denotes one of the two parts composing the linguistic sign. More specifically, the first constituent of the compound refers to the signifier, i.e. the sound-image (*signe*, *son*, *forme*), while the second element denotes the conceptual side, the idea conveyed by the sign (*idée*, *sens*).

Later, in a time span ranging from 1899 to 1903, Saussure renames *signe* as *sème* and distinguishes it from its "material" constituent, named *apossème* or *sôme* and from its conceptual part, designated as *parasôme* or *contre-sôme*, thus introducing a constellation of neologisms derived from ancient Greek, as is usual in scientific terminology. All these neologisms are abandoned and no longer adopted by the Genevan linguist who prefers to continue using more traditional terms such as *mot*, *terme*, *signe linguistique*. These terms are all attested in the third course of general linguistics held by Saussure in Geneva between 1910 and 1911. Here Saussure definitely chooses the term *signe* to indicate the whole resulting from the association of form, named *signifiant*, and meaning, named *signifié*, thus resolving the ambiguity that until then had characterized the term *signe*: "L'ambiguïté disparaîtrait si l'on désignait les trois notions ici en présence par des noms qui s'appellent les uns les autres tout en s'opposant. Nous proposons de conserver le mot *signe* pour désigner le total, et de remplacer concept et image acoustique respectivement par *signifié* et *signifiant*; ces derniers termes ont l'avantage de marquer l'opposition qui les sépare soit entre eux, soit du total dont ils font partie. [...]" [Saussure 1972, 99].

Another query focussing on the period of attestation of the term *signe* in all its meanings brings out this ambiguity more clearly, as illustrated in Figure 10.

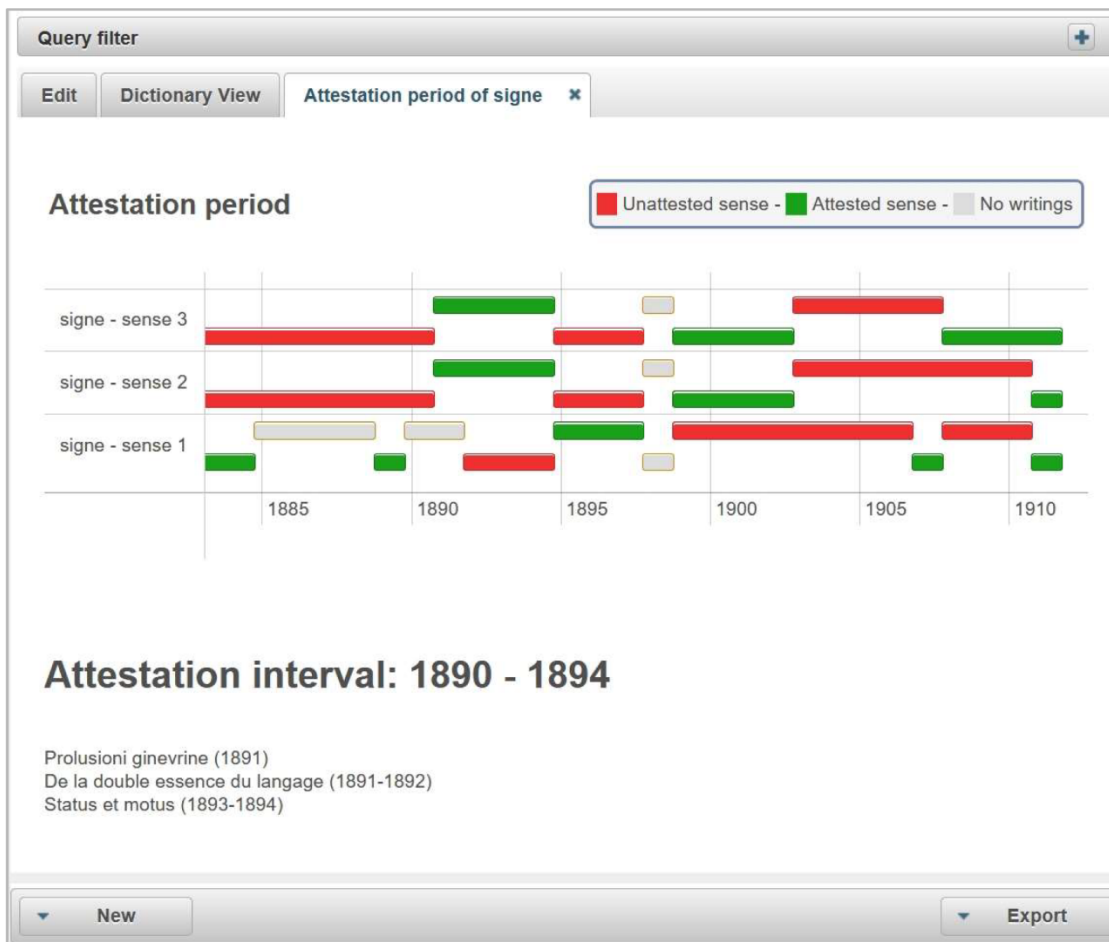


Figure 10. The attestation period of the term *signe* in its three meanings.

The green bar represents the period in which every sense is attested; clicking on it in LexO interface the complete list of all the works where the term occurs is provided.

As the graph shows, *signe* is used in the same period - and in the same texts - with two different meanings, namely to indicate the most “concrete” part of the linguistic sign (*signe sense-2*) as well as to denote the union of a concept and a form (*signe sense-3*).^[27] This polysemy generates ambiguity, and – as we know – ambiguity should be strongly avoided in terminology where, according to GTT, the term must be monosemic and the concept unambiguous. Saussure is aware of this and therefore during his last course on general linguistics – and more precisely on May 19 1911 – proposes to preserve the term *signe* in its third sense and to replace it with the neologism *signifiant* as for the second meaning.

As Saussure’s case study demonstrates, models and tools devoted to representing and querying diachronic terminological resources play a crucial role in order to better understand the terminological fluctuation and the mobile and progressive conceptualisation which characterize some unstable phases of science.

7. Conclusion and Next Steps

Representing the diachronic evolution of terminological and conceptual data in a formal way constitutes a challenge, since standard formalisms adopted in the Semantic Web for the representation of data have mainly been conceived for static and monotonic information.

In this work we have proposed a Semantic Web-based model for representing in an explicit and formal way the diachronic evolution of concepts and terms in a given domain, accompanied by a concrete use case focussing on the evolution of linguistic terminology as attested in Ferdinand De Saussure’s writings. The temporalization approach

illustrated here is domain independent and can be virtually applied to any RDF data model. The main drawback is that such an approach requires a significant design effort in modelling and encoding specific sets of temporal rules. In addition, the properties to be temporalized need to be already known at the modelling time. However, in slightly “controlled” domains like terminology and lexicography, our approach turns out to be effective to describe the historical development of a language, and its many temporal variations.

The next research objectives are: i) the enrichment of the model, ii) the exploitation of the resource, iii) the extension of the tool LexO to support the construction of diachronic resources [Bellandi et al. 2018].^[28]

As far as the enrichment of the model is concerned, we plan to design new SWRL rules for the automatic temporalization of lexico-semantic properties holding between senses, such as synonymy, antonymy etc.

In addition, in order to manage the cases when the writing period of a text is uncertain, we intend to develop an algorithm exploiting the resource for automatic dating suggestion: the behaviour of terms in those texts, whose writing time is known, will be used to advance hypotheses on the uncertain (or even unknown) writing time periods of the other texts belonging to the corpus under analysis. Needless to say, the accuracy of suggested time periods will be proportionate to the amount of (certain) temporal information already encoded in the diachronic resource.

Finally, we are working to extend the tool LexO as to include the management of texts and temporal information. The tool will be designed to relieve scholars from the time consuming (and prone to errors) task of formalising the large amount of entities required to represent the temporal information according to the N-ary model. In addition, thanks to LexO scholars will not be required to manage the formalisms and the languages underlying the Web Semantic technologies such as RDF and OWL.

The first extension of the tool LexO aimed at the inclusion of temporal information will involve the integration of functions for the management of text. As discussed in Section 2, diachronic terminology is a textual terminology, and terms have to be analysed in their contexts of use. For this reason, LexO will be equipped with a module through which users will be able to import their documents and the relative time information (writing time, date of publishing, etc.), organize them into folders, browse them and link each encoded term to the relative attestations in the corpus.

Notes

[1] The term terminology is here to be understood as the set of “lexical items belonging to specialized areas of usage of one or more languages” [Sager 1990, 2].

[2] “Clavius on the Web” [Piccini 2016], “Per un’edizione digitale dei manoscritti di Ferdinand de Saussure” [Ruimy et al. 2013], and “Totus Mundus” [Piccini 2017].

[3] <https://www.w3.org/RDF/>

[4] According to Gruber (1993: 199), “an ontology is a specification of a conceptualization”.

[5] https://en.wikipedia.org/wiki/Hypertext_Transfer_Protocol

[6] <https://lemon-model.net/> and <https://www.w3.org/2016/05/ontolex/> for details, see McCrae et al. (2017)

[7] Please refer to <https://www.w3.org/2016/05/ontolex/#core>

[8] <https://www.w3.org/TR/swbp-n-aryRelations/>

[9] <http://www.w3.org/TR/owl-time/>

[10] <https://www.w3.org/Submission/SWRL/>

[11] Rules having conjunctive consequents could easily be transformed into multiple rules each with an atomic consequent.

[12] For a complete built-ins list, please refer to Section 8 in <https://www.w3.org/Submission/SWRL/>.

94

95

96

97

98

[13] <https://www.w3.org/TR/swbp-n-aryRelations/>

[14] The plug-in for Protégé CHRONOS, devoted to creating and editing temporal ontologies in OWL, is based on this model.

[15] <http://purl.org/dc/dcmitype/Event>

[16] <https://www.w3.org/TR/owl-time/#time:ProperInterval>

[17] <http://purl.org/dc/terms/date>

[18] Please refer to <https://www.lexinfo.net/>

[19] In this paper, we considered exclusively these properties. Other lexico-semantic properties can be taken into account as well, for example translation defined in Lexinfo vocabulary (see <https://www.lexinfo.net/ontology/3.0/lexinfo>).

[20] <http://purl.org/dc/dcmitype/Text>

[21] <http://lawd.info/ontology/hasAttestation>

[22] It is important to underline that every relation defined by the scholar is - by default - considered valid for the entire time interval of the reference corpus.

[23] <http://purl.org/dc/terms/issued>

[24] <https://github.com/protegeproject/swrlapi/wiki/ExtensionsBuiltInLibrary>

[25] See Saussure's letter to Gaston Paris (30 December 1891). On the date of this manuscript, see Gambarara (2009).

[26] For the time being, LexO supports the uploading of diachronic resources, which are built according to the N-ary model, and offers a *Question Answering* system, implemented as a user-friendly interface composed of several controlled natural language templates. Each template makes it possible to formulate a query in controlled natural language thanks to dropdown menus. For further extensions of the tool LexO, see Section 7.

[27] *Signe sense-1* refers to any material entity to which we associate a sense. This meaning is not taken into account here, as not relevant for the purposes of this discussion.

[28] The most recent stable version of LexO is available at <https://github.com/andreabellandi/LexO-lite>.

Works Cited

- Allen 1983** Allen, J. (1983). "Maintaining knowledge about temporal intervals". *Communications of the ACM*, 26(11), pp. 832-843.
- Artale and Franconi 2000** Artale, A. and Franconi, E. (2000). "A survey of temporal extensions of description logics". *Annals of Mathematics and Artificial Intelligence*, 30 (1-4), pp. 171-210.
- Aráuz et al. 2012** Aráuz L., Faber, P. and Montero Martínez, S. (2012). "Specialized Language Semantics". In P. Faber, ed., *A cognitive linguistics view of terminology and specialized language*, Series: Applications of Cognitive Linguistics [ACL] 20, 1st ed. Berlin, Boston: De Gruyter Mouton, pp. 95-175.
- Baader and Sattler 2001** Baader, F. and Sattler U. (2001). "An overview of tableau algorithms for description logics". *Studia Logica*, 69 (1), pp. 5-40.
- Baader et al. 2003** Baader, F., Calvanese, D., McGuinness, D., Patel-Schneider, P., Nardi, D. (2003). "The description logic handbook: Theory, implementation and applications". Cambridge: Cambridge University Press.
- Batsakis et al. 2017** Batsakis, S., Petrakis E., Tachmazidis, I. and Antoniou, G. (2017). "Temporal representation and reasoning in OWL 2". *Semantic Web*, 8 (6), pp. 981-1000.
- Bellandi et al. 2017** Bellandi, A., Boschetti, F., Khan, F., Del Grosso, A.M. and Monachini, M. (2017). "Provando e riprovando modelli di dizionario storico digitale: Collegare voci, citazioni, interpretazioni". *Proceedings of the AIUCD*, pp. 119-125.

- Bellandi et al. 2018** Bellandi, A., Giovannetti, E., Weingart, A. (2018). "Multilingual and Multiword Phenomena in a lemon Old Occitan Medico-Botanical Lexicon". *Information* 9 (3), 52.
- Bellandi, Giovannetti, and Piccini 2018** Bellandi, A., Giovannetti, E., Piccini, S. (2018). "Collaborative Editing of Lexical and Terminological resources: A Quick Introduction to LexO". *The XVIII EURALEX International Congress: Lexicography in Global Contexts Book of Abstracts*, 17-21 July 2018, Ljubljana, Slovenia, pp. 23-27.
- Benveniste 1964** Benveniste, E. (1964). "Lettres de F. de Saussure à A. Meillet". *Cahiers Ferdinand de Saussure*, 21, pp. 89-130.
- Berners-Lee et al. 2001** Berners-Lee, T., Hendler, J., Lassila, O. (2001). "The semantic web". *Scientific American*, 284(5), pp. 34-43.
- Bizer et al. 2011** Bizer, Ch., Heath, T., Berners-Lee, T. (2011). "Linked data: The story so far". In Amit Sheth, ed., *Semantic services, interoperability and web applications: emerging concepts*, 1st ed., Hershey, Pennsylvania: IGI Global, pp. 205-227.
- Bourigault and Slodzian 1999** Bourigault D., Slodzian M. (1999). "Pour une terminologie textuelle". *Terminologies nouvelles* 19, pp. 29-32.
- Buitelaar 2010** Buitelaar, P. (2010). "Ontology-based Semantic Lexicons: Mapping between Terms and Object Descriptions", In C. Huang, N. Calzolari, A. Gangemi, A. Lenci, A. Oltramari, & L. Prevot, eds., *Ontology and the Lexicon: A Natural Language Processing Perspective. Studies in Natural Language Processing*. Cambridge: Cambridge University Press, pp. 212-223.
- Cabré 1999** Cabré, M. T. (1999). *Terminology: Theory, Methods and Applications*. Amsterdam: John Benjamins Publishing.
- Calvanese et al. 2001** Calvanese, D., De Giacomo, G., Lenzerini, M. Nardi, D. (2001). "Reasoning in Expressive Description Logics". *Handbook of automated reasoning* 2, pp. 1581-1634.
- Carlyle 2011** Carlyle, A. (2011). "Understanding FRBR as a conceptual model". *Library Resources & Technical Services*, 50 (4), pp. 264-273.
- Cimiano et al. 2011** Cimiano, P., Buitelaar, P., McCrae, J., Sintek, M. (2011). "Lexinfo: A declarative model for the lexicon-ontology interface". *Web Semantics: Science, Services and Agents on the World Wide Web* 9 (1), pp. 29-51.
- Cimiano et al. 2015** Cimiano, P., McCrae, J.P., Rodríguez-Doncel, V., Gornostay, T., Gómez-Pérez, A., Siemoneit, B., Lagzdins, A. (2015). "Linked terminology: applying linked data principles to terminological resources". In Kosem, I., Jakubiček, M., Kallas, J., Krek, S., eds., *Electronic lexicography in the 21st century: linking lexical data in the digital age. Proceedings of the eLex 2015 conference*, 11-13 August 2015, Herstmonceux Castle, United Kingdom. Ljubljana/Brighton: Trojina, Institute for Applied Slovene Studies/Lexical Computing Ltd, Sussex, UK, pp. 504-517.
- Cosenza 2016** Cosenza, G. (2016). *Dalle parole ai termini. I percorsi di pensiero di F. de Saussure*. Alessandria. Edizioni dell'Orso. Collezione "Studi e Ricerche".
- Dechandon et al. 2019** Dechandon, D., Gerencsér, A., Ruiz, M.R. (2019). "Terminology: Towards a Systematic Integration of Semantics and Metadata". In *Translating and the Computer* 41, Proceedings of AsLing (The International Association for Advancement in Language Technology), Geneva: Editions Tradulex.
- Del Gratta et al. 2015** Del Gratta, R., Frontini, F., Khan, F., Monachini, M. (2015). "Converting the PAROLE SIMPLE CLIPS Lexicon into RDF with lemon". *Semantic Web*, 6 (4), pp. 387-392.
- Dury 2004** Dury, P. (2004). "Building a Bilingual Diachronic Corpus of Ecology: The Long Road to Completion". *Icame Journal*, 28, pp. 5-16.
- Dury 2008** Dury, P. (2008). "The rise of carbon neutral and compensation carbone: A diachronic investigation into the migration of vocabulary from the language of ecology to newspaper language and vice versa". *Terminology*, 14(2), pp. 230-248.
- Dury and Picton 2009** Dury, P., Picton, A. (2009). "Terminologie et diachronie: vers une réconciliation théorique et méthodologique?". *Revue française de linguistique appliquée*, 14 (2), pp. 31-41.
- Eckle-Kohler 2015** Eckle-Kohler, J., McCrae, J.P., Chiarcos, Ch. (2015). "LemonUby—A large, interlinked, syntactically-rich lexical resource for ontologies". *Semantic Web*, 6 (4), pp. 371-378.
- Ehrmann et al. 2014** Ehrmann, M., Cecconi, F., Vannella, D., McCrae, J.P., Cimiano, P., Navigli, R. (2014). "Representing Multilingual Data as Linked Data: the Case of BabelNet 2.0". In N. Calzolari, K. Choukri, T. Declerck, H. Loftsson, B. Maegaard, J. Mariani, A. Moreno, J. Odijk, S. Piperidis, eds., *Proceedings of LREC 2014*, pp. 401-408.

- Engler 1968** Engler, R. (1968). *Lexique de la terminologie saussurienne*. Utrecht-Anvers: Spectrum. Comité international permanent des linguistes. Publication de la commission de terminologie.
- Fenoglio 2012** Fenoglio, I. (2012). "Genèse du geste linguistique : une complexité heuristique". *Genesis* [En ligne], 35, pp. 13-40. DOI : 10.4000/genesis.1033.
- Flouris et al. 2008** Flouris, G., Manakanatas D., Kondylakis H., Plexousakis D., Antoniou G. (2008). "Ontology Change: Classification and Survey". *The Knowledge Engineering Review*, 23, pp. 117-52.
- Francopoulo et al. 2006** Francopoulo, G., George, M., Calzolari, N., Monachini, M., Bel, N., Pet, M., Soria, C. (2006). "Lexical markup framework (LMF)". *Proceedings of the Fifth International Conference on Language Resources and Evaluation LREC 2006*, Genova, pp. 233-236.
- Gambarara 2009** Gambarara, D. (2009). Présentation de section II "Système et cognition. Quaternion et parallélie dans De l'Essence double du langage". *Cahiers Ferdinand de Saussure*, 61, pp. 75-86.
- Godel 1957** Godel, R. (1957). *Les sources manuscrites du Cours de linguistique générale de Ferdinand de Saussure*. Genève: Droz.
- Grandi 2009** Grandi, F. (2009). "Multi-temporal RDF ontology versioning". In M. d'Aquin, G. Antoniou, eds., *Proceedings of the 3rd International Workshop on Ontology Dynamics (IWOD-09)*, collocated with the 8th International Semantic Web Conference (ISWC-2009), Washington DC, USA, October 26, 2009.
- Gruber 1991** Gruber, T. R. (1991). "A translation approach to portable ontologies". *Knowledge Acquisition*, vol. V (2), pp. 199-220.
- Kanmani 2016** Kanmani, A. Clara, Chockalingam, T. and Guruprasad. N. (2016). "RDF data model and its multi reification approaches: A comprehensive comparative analysis". In *International Conference on Inventive Computation Technologies (ICICT)*, vol. 1, pp. 1-5.
- Khan et al. 2014** Khan, F., Boschetti, F., Frontini, F. (2014). "Using lemon to model lexical semantic shift in diachronic lexical resources". In Ch. Chiarcos, J. Ph. McCrae, P. Osenova, C. Vertan, eds., *3rd. Workshop on Linked Data in Linguistics: Multilingual Knowledge Resources and Natural Language Processing*, p. 50.
- Khan et al. 2016** Khan, F., Bellandi, A., Monachini, M. (2016). "Tools and Instruments for Building and Querying Diachronic Computational Lexica". *Proceedings of the Workshop on Language Technology Resources and Tools for Digital Humanities (LT4DH)*, Osaka, Japan: The COLING 2016 Organizing Committee, pp. 164-171.
- Khan et al. 2017** Khan, F., Bowers, J., Frontini, F. (2017). "Situating Word Senses in their Historical Context with Linked Data". In C. Gardent, Ch. Retoré, eds., *IWCS 2017 — 12th International Conference on Computational Semantics — Short papers*.
- Klein and Fensel 2001** Klein, M. CA., Fensel, D. (2001). "Ontology versioning on the Semantic Web". *Proceedings of the First International Conference on Semantic Web Working*. California: CEUR-WS.org, pp. 75-91.
- Kripke 1995** Kripke, S. (1995). *La logique des noms propres*. Paris: Les Editions de Minuit.
- Kristiansen 2014** Kristiansen M. (2014). "Concept change, term dynamics and culture-boundness in economic-administrative domains". In R. Temmermann and M. van Campenhoudt, eds., *Dynamics and terminology: an interdisciplinary perspective on monolingual and multilingual culture-bound communication*. Terminology and lexicography research and practice, volume 16, Amsterdam; Philadelphia: John Benjamins Publishing Company, pp. 235-256.
- Kroeger 2013** Kroeger, A. (2013). "The road to BIBFRAME: the evolution of the idea of bibliographic transition into a post-MARC Future". *Cataloging & classification quarterly*, 51 (8), pp. 873-890.
- Kuhn 1970** Kuhn, T.S. (1970). *The structure of scientific revolutions*. 2nd ed., Chicago: The University of Chicago Press.
- Lacy 2005** Lacy, L.W. (2005). *Owl: Representing Information Using the Web Ontology Language*. Victoria, BC, Canada: Trafford Publishing.
- Lenci et al. 2000** Lenci, A., Bel, N., Busa, F., Calzolari, N., Gola, E., Monachini, M., Ogonowski, A., Peters, I., Peters, W., Ruimy, N., Villegas, M., Zampolli, A. (2000). "SIMPLE: A General Framework for the development of Multilingual Lexicons". *International Journal of Lexicography*, 13 (4). Special issue "Dictionaries, Thesauri and Lexical-Semantic Relations", pp. 249-263.
- Lutz et al. 2008** Lutz, C., Wolter, F., Zakharyashev, M. (2008). "Temporal Description Logics: A Survey". *TIME*, pp. 3-14.

- Makri 2011** Makri, P. (2011). 4D-Fluents Plug-In: A Tool for Handling Temporal Ontologies in Protégé. *Diss. Diploma Thesis*. Department of Electronic and Computer Engineering, Technical University of Crete.
- McCrae et al. 2014** McCrae, J.P., Fellbaum, C., Cimiano, P. (2014). "Publishing and linking WordNet using lemon and RDF". *Proceedings of the 3rd Workshop on Linked Data in Linguistics*.
- McCrae et al. 2017** McCrae, J. P., Bosque-Gil, J., Gracia, J., Buitelaar, P., & Cimiano, P. "The OntoLex-Lemon Model: Development and Applications". In I. Kosem, C. Tiberius, M. Jakubíček, J. Kallas, S. Krek, V. Baisa, eds, *Electronic lexicography in the 21st century. Proceedings of eLex 2017 conference*, September 19-21, Leiden, Netherlands. Lexical Computing CZ s.R.O, Brno, Czech Republic, pp. 19-21.
- Montiel-Ponsoda 2008** Montiel-Ponsoda, E., Aguado de Cea, G., Gómez-Pérez, A., Peters, W. (2008). "Modelling multilinguality in ontologies". *Proceedings of the 21st International Conference on Computational Linguistics (COLING)*.
- Møller 1998** Møller, B. (1998). "À la recherche d'une terminochronie". *Meta*, XLIII-3, pp. 1-13
- Pearson 1998** Pearson, J. (1998). *Terms in Context, Studies in Corpus Linguistic*. Vol. I, 1 ed., Amsterdam/Philadelphia: John Benjamins.
- Piccini 2016** Piccini, S. (2016). "CLAVIUS: verso la modellazione di una risorsa termino-ontologica diacronica del dominio matematico-astronomico del XVII secolo". *Atti del XXVI Convegno internazionale Ass.I. Term. Terminologia e organizzazione della conoscenza nella conservazione della memoria digitale*, 14 - 16 April 2016 - Rende, Italy, pp. 95-105.
- Piccini 2017** Piccini, S., Corsi, E., Giovannetti, E. (2017). "Une ressource termino-ontologique bilingue chinois-italien: le cas de la traduction de la Mappemonde de Matteo Ricci par Pasquale D'Elia". In Ch. Roche, éd., *TOTh 2017, Terminologie & Ontologie : Théories et Applications*, Série: Terminologica, Éditions de l'Université Savoie Mont Blanc, pp. 33-49.
- Piccini 2018** Piccini, S., Bellandi, A., Giovannetti, E. (2018). "A Semantic Web Approach to Modelling and Building a Bilingual Chinese-Italian Termino-ontological Resource". In A. Čibej, V. Gorjanc, I. Kosem, S. Krek, eds., *The XVIII EURALEX International Congress: Lexicography in Global Contexts Book of Abstracts*, 17-21 July 2018, Ljubljana, Slovenia, pp. 87-90.
- Picton 2009** Picton, A. (2009). *Diachronie en langue de spécialité. Définition d'une méthode linguistique outillée pour repérer l'évolution des connaissances en corpus. Un exemple appliqué au domaine spatial*. PhD. Dissertation. Toulouse 2.
- Preventis et al. 2014** Preventis, A., Petrakis, E., Batsakis, S. (2014). "Chronos Ed: A tool for handling temporal ontologies in Protégé". *International Journal on Artificial Intelligence Tools*, 23 (4).
- Pustejovsky 1995** Pustejovsky, J. (1995). *The Generative Lexicon*, 1st ed. Cambridge MA: The MIT Press.
- Rouces et al. 2017** Rouces, J., de Melo, G., Hose, K. (2017). "FrameBase: Enabling integration of heterogeneous knowledge". *Semantic Web*, 8(6), pp. 817-850.
- Ruimy et al. 2013** Ruimy, N., Piccini, S., Giovannetti, E., Bellandi, A. (2013). "Lessicografia Computazionale e Terminologia Saussuriana". In D. Gambarara, M.P. Marchese, eds., *Guida per un'edizione digitale dei manoscritti di Ferdinand de Saussure*, Alessandria: Edizioni dell'Orso.
- Sager 1990** Sager J. (1990). *A Practical Course in Terminology Processing*. Amsterdam: John Benjamins Publishing.
- Salvadores et al. 2013** Salvadores, M., Alexander, P.R., Musen, M.A., Noy, N.F. (2013). "Bioportal as a dataset of linked biomedical ontologies and terminologies in RDF". *Semantic Web Journal*, 4(3), pp. 277-284.
- Sansone 2007** Sansone, S. A., Scheuermann, R. H., Shah, N., Whetzel, P. L., Lewis, S. (2007). "The OBO Foundry: coordinated evolution of ontologies to support biomedical data integration". *Nature biotechnology*, 25(11), 2007:1251.
- Saussure 1972** Saussure de, F. (1972). *Cours de linguistique générale*. Edition critique préparée par Tullio De Mauro, Paris: Payot.
- Slodzian 2000** Slodzian M. (2000). "L'émergence d'une terminologie textuelle et le retour du sens". In H. Béjoint and Ph. Thoiron, eds., *Le sens en terminologie*, Lyon: Presse universitaire de Lyon, pp. 61-85.
- Smith et al. 2007** Smith, B., Ashburner, M., Rosse, C., Bard, J., Bug, W., Ceusters, W., Goldberg, L.J., Eilbeck, K., Ireland, A., Mungall, C.J.; OBI Consortium, Leontis, N., Rocca-Serra, P., Ruttenberg, (2007) "The OBO Foundry: coordinated evolution of ontologies to support biomedical data integration". *Nature Biotechnology*, 25, pp. 1251-1255.

- Stetter 1979** Stetter, C. (1979). "La fonction des réflexions sémiologiques dans la fondation de la linguistique générale chez F. de Saussure". *Kodikas-Code*, Gunter Narr Verlag, Tübingen, vol. 1, n° I, p. 12.
- Sánchez Ibáñez 2014** Sánchez Ibáñez, M., García Palacios, J. (2014). "Semantic characterization of terms as a trace of terminological dependency". *Terminology. International Journal of Theoretical and Applied Issues in Specialized Communication*, 20 (2), pp. 171-197.
- Sánchez-Rada et al. 2014** Sánchez-Rada, J. F., Vulcu, G., Iglesias C. A., and Buitelaar, P. (2014). "EUROSENTIMENT: Linked data sentiment analysis". In M. Horridge, M. Rospocher, J. van Ossenbruggen, eds., *Proceedings of the ISWC 2014 Posters & Demonstrations Track, a track within the 13th International Semantic Web Conference (ISWC 2014)*, pp. 145-148.
- Sérasset 2015** Sérasset, G. (2015). "DBnary: Wiktionary as a Lemon-based multilingual lexical resource in RDF". *Semantic Web*, 6 (4), pp. 355-361.
- Tappolet and Bernstein 2009** Tappolet, J., Bernstein, A. (2009). "Applied temporal RDF: Efficient temporal querying of RDF data with SPARQL". *European Semantic Web Conference*, Berlin, Heidelberg: Springer, pp. 308-322.
- Van Campenhoudt 1997** Van Campenhoudt, M. (1997). "Maille ou maillon : quand des terminographes négligent l'évolution de l'usage". In A. Clas, S. Mejri, T. Baccouche, eds., *Proceedings of the 5th scientific days Réseau: Lexicologie, Terminologie, Traduction* (Agence Universitaire de la Francophonie), *La mémoire des mots*, Tunis, pp. 251-272.
- Welty et al. 2006** Welty, Ch., Fikes, R., Makarios, S. (2006). "A reusable ontology for fluents in OWL". *FOIS*, vol. 150, pp. 226-236.
- Wilkinson et al. 2016** Wilkinson, M.D., Dumontier, M., Aalbersberg, I.J., Appleton, G., Axton, M., Baak, A., Blomberg N. et al. (2016). "The FAIR Guiding Principles for scientific data management and stewardship". *Scientific data* 3.
- Wüster 1968** Wüster, E. (1968). *The Machine Tool. An Interlingual dictionary of basic concepts*. 1 ed., London: Technical Press.
- Zanola 2014** Zanola, M. T. (2014). *Arts et métiers au XVIIIe siècle. Essais de terminologie diachronique*. Paris: L'Harmattan.
- Zekri et al. 2017** Zekri A., Brahmia, Z., Grandi, F., Bouaziz, R. (2017). "Temporal schema versioning in τ OWL: a systematic approach for the management of time-varying knowledge". *Journal of Decision Systems*, 26 (2), pp. 113-137.