Tools and Instruments for Building and Querying Diachronic Computational Lexica

Anas Fahad Khan[*][†]
[*]Dipartimento di Studi Umanistici
Universita' di Ca' Foscari
Dorsoduro 3246, 30123 Venezia, Italy
fahad.khan@unive.it

Andrea Bellandi[†], **Monica Monachini**[†]
[†]Istituto di Linguistica Computazionale
"Antonio Zampolli"
Via G. Moruzzi 1, 56121 Pisa, Italy
name.surname@ilc.cnr.it

Abstract

This article describes work on enabling the addition of temporal information to senses of words in linguistic linked open data lexica based on the *lemonDia* model. Our contribution in this article is twofold. On the one hand, we demonstrate how *lemonDia* enables the querying of diachronic lexical datasets using OWL-oriented Semantic Web based technologies. On the other hand, we present a preliminary version of an interactive interface intended to help users in creating lexical datasets that model meaning change over time.

1 Introduction

The necessity for a flexible and accurate way of representing diachronic lexical information computationally is particularly evident when dealing with "classical" languages such as Ancient Greek, Latin or Sanskrit where we have access to a corpus of texts covering a long period in the language's evolution. It is also the case for modern languages like French, Italian or English where we can count on an existing legacy of texts that attest to various different periods in those languages' development. An important requirement for representation formats for diachronic lexico-semantic resources is that they should facilitate cross-linguistic, typological research of the kind that takes into account different language features both across diverse languages as well as different time periods. One way of working towards meeting such requisites is through the adoption of the linked open data (LOD) paradigm as a means of modelling and publishing such lexical datasets. This not only ensures a minimum level of inter-operability between different datasets through the shared use of the Resource Data Framework (RDF) and common vocabularies/data categories, but it also allows us to exploit various RDF based technologies such as the Web Ontology Language (OWL) when working with such data. We discuss this in more detail below.

In this article we will focus on a model/vocabulary for representing diachronic semantic information as RDF triples called *lemonDia*, which we have introduced in previous work (see (Khan et al., 2014) and (Khan et al., 2016)). In the present work we will look at the more practical aspects of using *lemonDia* and show how *lemonDia* enables the querying of diachronic lexical datasets, using two OWL oriented Semantic Web based technologies, the Semantic Web Rule Language (SWRL) and the Semantic Query-Enhanced Web Rule Language (SQWRL). We will also introduce an interactive tool which we are developing and which is intended to assist users in creating lexical linked open datasets that include information about meaning change over time. One of the main difficulties with incorporating temporal information within RDF datasets is that the rigid subject-predicate-object triple structure of RDF prevents the addition of an extra time argument; this can be resolved in several ways, none of which are entirely satisfactory. *lemonDia* uses the modelling 'trick' of explicitly representing senses as processes in time, or perdurants, but this can be difficult to grasp for non-expert users. Making the whole process of working of assigning temporal periods to lexical entries easier and therefore making the *lemonDia* model more accessible, was one of the main motivations behind the creation of our interactive tool.

Proceedings of the Workshop on Language Technology Resources and Tools for Digital Humanities (LT4DH), pages 164–171, Osaka, Japan, December 11-17 2016.

The article is organized as follows: Section 2 presents the diachronic dataset that we worked with and that provided the main case study for our tool; Section 3 briefly describes *lemonDia* and Section 4 shows how it is possible to make temporal queries over the dataset from Section 2; Section 5 describes the preliminary version of our interactive tool. Finally, Section 6 draws the conclusions and outlines future work.

2 The Old English Shame/Guilt Dataset

The examples which we will present in this article are taken from a lexical dataset of Old English (OE) emotion terms produced by Díaz-Vera (E Díaz-Vera, 2014) as part of a wider study into the cognitive implications of meaning change. The lexical entries in the dataset have been categorized into those relating to shame/embarrassment and those relating to guilt. The dataset contains both emotion terms which are classified as "literal" – that is emotion terms that aren't the result of a semantic shift from another domain – and non-literal terms, where there is a clear shift from another domain into the domain of shame or guilt. These latter are classified further on the basis of the semantic shifts in question. The time period in which Old English was spoken is divided into 3 consecutive intervals in the OE dataset. These are:

- OE1 (before 950)
- OE2 (950-1050)
- OE3 (1050-1150) .

For simplicity the literal word senses in the dataset are assumed to be valid throughout the whole period in which Old English was spoken. Other senses have an associated period which corresponds to one or more of the three individual periods. These periods can be encoded in RDF-OWL as proper intervals using ProperInterval from the time vocabulary[1]. We have encoded the second interval, OE2, as follows:

```
:OE2 rdf:type owl:NamedIndividual ,
    <http://www.w3.org/2006/time#ProperInterval> ;
    <http://www.w3.org/2006/time#hasBeginning> :year_950
    <http://www.w3.org/2006/time#hasEnd> :year_1050 .
```

3 Using *lemon* and *lemonDIA*

lemon was originally intended as a model for enriching ontologies with linguistic information (McCrae et al., 2010). However it quickly came to take on the status of a de facto standard for representing lexicons as linked open data. Indeed *lemon* has so far been used to convert the Princeton WordNet and Wiktionary (McCrae et al., 2012), as well as FrameNet and VerbNet (Eckle-Kohler et al., 2015), among other well known resources. The design of *lemon* was heavily influenced by the Lexical Markup Framework (LMF) (Francopoulo et al., 2006), but with numerous simplifications to the original LMF specifications. In addition unlike LMF the *lemon* model focuses specifically on creating lexico-semantic resources with an ontological component where the ontology represents the extensions of the word senses in the lexicon. So that in the *lemon* model every lexical sense necessarily connects a lexical entry with a specific ontology vocabulary item. *lemonDia* (Khan et al., 2014) was designed as an extension for *lemon* with the specific purpose of enabling the addition of temporal information to senses. We felt this was necessary even though the original *lemon* model did have a usedSince property (a subproperty of lemon:context) which allowed users to specify the date from which a word was used with a given sense. However this property by itself was clearly not flexible enough to represent the evolution of word senses over time.

The main idea with *lemonDia* is to add a temporal parameter to the sense relation linking together a Lexical Entry and a Lexical Sense. As mentioned above RDF has the restriction that all statements must conform to a subject-predicate-object structure. We therefore chose to treat lexical senses as *perdurants*

[1] http://www.w3.org/2006/time

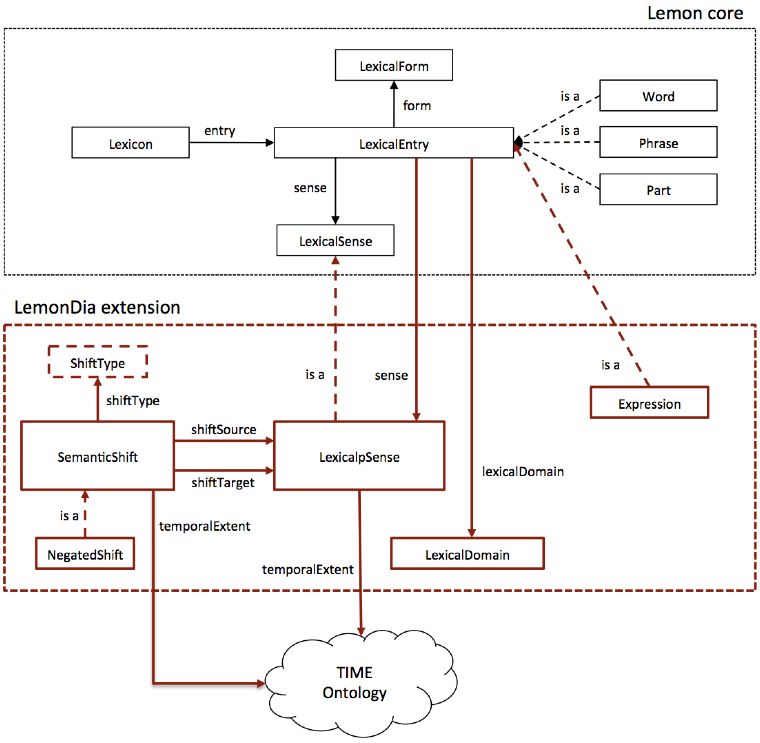


Figure 1: The lemonDia model (LexicalSense and LexicalpSense refer to ontology elements).

with an inherent temporal extent and defined a special subclass of LexicalSense in *lemonDia* called LexicalpSense. Every member of LexicalpSense has an associated time period, a time:ProperInterval which it is linked to via the *lemonDia* temporalExtent property. A graphical representation of the *lemonDia* extension is depicted in Figure 1, and more details are given in (Khan et al., 2016). So for instance the lexical entry for the verb *areodian* from the OE dataset, meaning both ‘to turn red’ as well as the more specific ‘to redden with shame’ can be linked to two different instances of LexicalpSense corresponding to each of these two senses:

```
:AREODIAN_VB a lemon:LexicalEntry ;
    lemon:language "ang" ;
    lemon:sense :sense_Red_AREODIAN_VB,
    :sense_Shame_AREODIAN_VB ;
    lexinfo:partOfSpeech lexinfo:Verb .
```

Since these two senses are perdurants they have an associated temporal extent. In the next section we will look at how to use *lemonDia* and other semantic web technologies to encode and then query, information about these senses and their associated temporal intervals.

4 Entering and Querying Temporal Information

When it comes to working with temporal intervals representing the time periods in which a given language (or language variety) was spoken, or in which a given word had a specific sense, we have to reckon with the fact that in many cases we don’t have a specific start date – either in terms of a year or maybe even a century – or, when appropriate, an end date. This is of course presents a major obstacle in querying such datasets using a language such as SPARQL. Fortunately, we can overcome this lack of data through the use of Allen’s basic relations to define time periods in terms of their relations to each other (Allen, 1983), that is qualitatively instead of quantitatively (see (Batsakis et al., 2009)). For instance we can define the time period in which proto-European was spoken in terms of the fact that it came before the period in which proto-Germanic was spoken, which in its own turn gave birth to the Germanic languages for which we have written evidence.

An extremely useful resource for working with such time periods in OWL was developed by Batsakis and consists of a set of rules in the Semantic Web Rule Language (SWRL) encoding the Allen relations[2]. SWRL is, as the name suggests, a rule language specifically designed for the semantic web; it is a subset of Datalog[3] and is strictly more expressive than OWL. Although SWRL as a whole is undecidable there is a subset of the language, the set of DL-safe rules, in which all variables appearing in the consequent of the rule must also appearin the antecedent, that is decidable. The rules and the examples that we are working with belong to this subset. Another Semantic Web technology that is relevant here is the Semantic Query-Enhanced Web Rule Language (SQWRL), a query language that allows the querying of OWL datasets using a query syntax based on SWRL. SQWRL is specifically tailored to querying datasets in OWL a task for which SPARQL is arguably less well adapted (O'Connor and Das, 2009).

SWRL rules, such as those developed by Batsakis, enable us to combine the basic intervals defined above, i.e., OE1, OE2, OE3, to define new intervals. For instance the temporal extent associated with the sense *sense_Shame_AREODIAN_VB*, OE23, is the sum of the two periods OE2 and OE3.

```
:sense_Shame_AREODIAN_VB a lemond:LexicalpSense ;
    lemon:reference dbpedia:Shame ;
    lemond:temporalExtent anglo:OE23 .
```

This period OE23 can be defined as follows, using the intervalStarts and intervalFinishes Allen relations and the two intervals OE2 and OE3 previously defined.

```
:OE23  rdf:type owl:NamedIndividual ,
            <http://www.w3.org/2006/time#ProperInterval> ;
            <http://www.w3.org/2006/time#intervalStarts> :OE2 ;
            <http://www.w3.org/2006/time#intervalFinishes> :OE3.
```

Using SQWRL we can write queries that exploit the logical axioms and rules in our dataset and that, using an OWL reasoner, are able to take into consideration knowledge, and in our case temporal knowledge, that is only implicit in the dataset itself. We now give three examples of queries typifying useful kinds of query that one can make on such a dataset. To start off with the following query will produce a list of all the lexical entries in the dataset and the number of senses which they have:

```
lemon:sense(?x, ?y) -> sqwrl:select(?x) ˆ sqwrl:count(?y)
```

We can also produce a list of all the senses that have a temporal extent of OE1:

```
lemond:LexicalpSense(?x) ˆ lemond:temporalExtent(?x,anglo: OE1)
-> sqwrl:select(?x)
```

The following query finds all the senses that contain the sense OE1:

```
lemond:LexicalpSense(?x) ˆ lemond:temporalExtent(?x,?y)
ˆintervalContains(?y, anglo:OE2)-> sqwrl:select(?x)
```

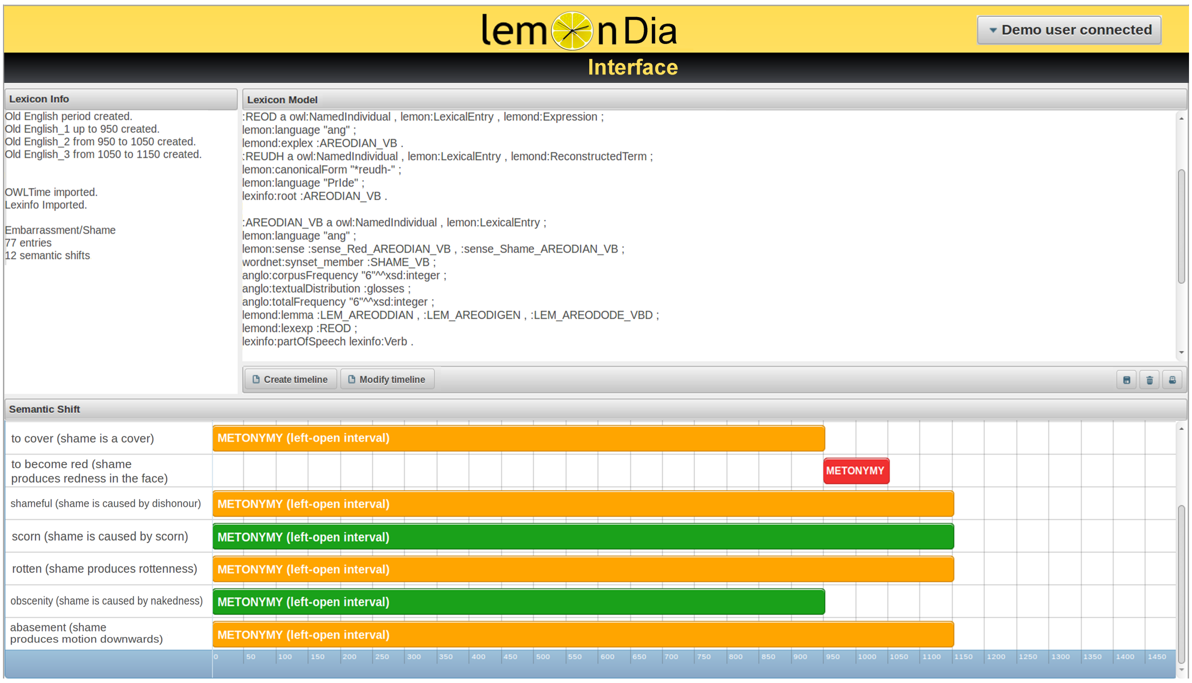
The use of SWRL and SQWRL seems to be gathering traction. The latest version of *Protégé* (*Protégé* 5.0.0), probably the most popular free tool available for constructing and editing ontologies, comes prepackaged with a tab for carrying out SQWRL queries on OWL datasets.

5 Our Interactive System for Creating *lemonDIA* lexicons

In the previous section we looked at some of the queries that can be performed on a dataset like the Old English one which we introduced earlier in the article. One of the motivations for this was to show the potential benefits of creating datasets using the *lemonDia* model. In order to overcome the initial hurdles to creating *lemonDia* datasets in the first place however, we have developed an interface which assists the user in doing this. Before we go on to describe this interface we will look in the following section at some related work.

[2] https://github.com/sbatsakis/TemporalRepresentations

[3] https://www.w3.org/Submission/SWRL/



lemonDia
Interface
Demo user connected
Lexicon Info
Lexicon Model
Create timeline
Modify timeline
Semantic Shift
to cover (shame is a cover)
METONYMY (left-open interval)
to become red (shame produces redness in the face)
METONYMY
shameful (shame is caused by dishonour)
METONYMY (left-open interval)
scorn (shame is caused by scorn)
METONYMY (left-open interval)
rotten (shame produces rottenness)
METONYMY (left-open interval)
obscenity (shame is caused by nakedness)
METONYMY (left-open interval)
abasement (shame produces motion downwards)
METONYMY (left-open interval)
0 50 100 150 200 250 300 350 400 450 500 550 600 650 700 750 800 850 900 950 1000 1050 1100 1150 1200 1250 1300 1350 1400 1450

Old English period created.
Old English_1 up to 950 created.
Old English_2 from 950 to 1050 created.
Old English_3 from 1050 to 1150 created.

OWLTime imported.
Lexinfo Imported.

Embarrassment/Shame
77 entries
12 semantic shifts

```
:REOD a owl:NamedIndividual , lemon:LexicalEntry , lemond:Expression ;
lemon:language "ang" ;
lemond:explex :AREODIAN_VB .
:REUDH a owl:NamedIndividual , lemon:LexicalEntry , lemond:ReconstructedTerm ;
lemon:canonicalForm "*reudh-" ;
lemon:language "PrIde" ;
lexinfo:root :AREODIAN_VB .

:AREODIAN_VB a owl:NamedIndividual , lemon:LexicalEntry ;
lemon:language "ang" ;
lemon:sense :sense_Red_AREODIAN_VB , :sense_Shame_AREODIAN_VB ;
wordnet:synset_member :SHAME_VB ;
anglo:corpusFrequency "6"^^xsd:integer ;
anglo:textualDistribution :glosses ;
anglo:totalFrequency "6"^^xsd:integer ;
lemond:lemma :LEM_AREODDIAN , :LEM_AREODIGEN , :LEM_AREODODE_VBD ;
lemond:lexexp :REOD ;
lexinfo:partOfSpeech lexinfo:Verb .
```

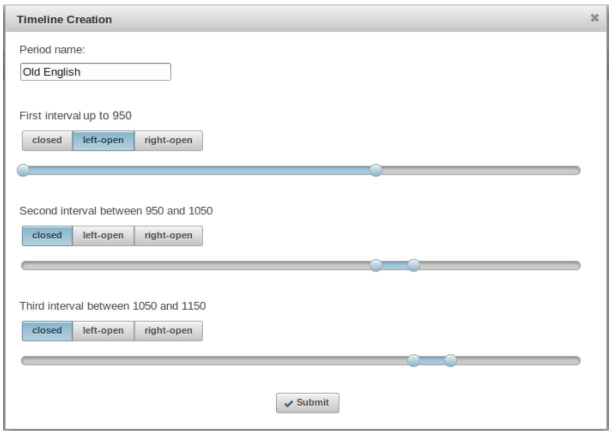
Figure 2: Main interface for creating and visualizing *lemonDia* lexica.

5.1 Related Work

A number of tools have been developed in order to work with standards and vocabularies such as LEXINFO (Cimiano et al., 2011), LMF, and *lemon*. For example, in (Johnson et al., 2005) the authors discuss LexGrid, a tool enabling the integration of terminologies and ontologies through a common model. They present an overview of the editor's functional capabilities in relation to technologies offered by the LexGrid platform.(Ringersma and Kemps-Snijders, 2007) describes the development of a flexible web based lexicon tool, LEXUS which allows the creation of lexica within the structure of the ISO LMF standard and uses the proposed concept naming conventions from the ISO data categories, thus enabling interoperability, search and merging. Another generic platform for working with computational lexica, is presented in (Bel et al., 2008): the COLDIC system has been specially designed to allow the user to concentrate on the lexicographical task at hand while being autonomous in the management of the tools. Montiel et. al. (Montiel-Ponsoda et al., 2008) propose a tool, developed as a plug-in of NeOn[4] to support a model called the Linguistic Information Repository (LIR). LIR is a holistic linguistic information repository, that provides a complete set of linguistic elements in each language for localizing ontology elements. It also allows access to linguistic information distributed in heterogeneous resources of varying granularities, and makes it possible to establish relations between linguistic elements. Another plug-in for the NeOn toolkit has been developed in (Buitelaar et al., 2009). (Touhami et al., 2011) proposes a new model whose use is illustrated within a supervised annotation environment in which the user can manually enrich an Ontological and Terminological Resource (OTR) by associating each new found term to the appropriate domain concepts. They have developed their OTR editor called TextViz as a plug-in in the *Protégé*-OWL framework. It also helps the user to visualize the textual manifestations of concepts in the corpus used to construct the OTR. Finally, in (Kenter et al., 2012), an editor for constructing corpus-based lexica and correcting word-level annotations and transcription errors in corpora, is presented. The editor has been extensively tested in a project in which a historical corpus was manually annotated and used to produce a lexicon, with the lexicon being further extended on the basis of a much larger corpus.

As regards *lemon*, in (McCrae and Unger, 2014) the authors use ontology design patterns (Gangemi, 2005) for defining how certain lexico-semantic phenomena should be modelled. Their goal in creating such a catalogue of ontology-lexicon design patterns is to facilitate the process of developing ontology-

[4] NeOn toolkit is available at `http://neon-toolkit.org`



Timeline Creation

Period name: Old English

First interval up to 950

closed [✓ left-open] right-open

Second interval between 950 and 1050

[✓ closed] left-open right-open

Third interval between 1050 and 1150

[✓ closed] left-open right-open

Figure 3: Interface for creating time intervals.

lexica, by replacing complex combinations of frame semantics and first-order logic axioms with simple patterns with only a few parameters. Finally, in (Montiel-Ponsoda et al., 2012), a platform called *lemon* source is presented. It supports the creation of linked lexical data and it builds on the concept of a semantic wiki to enable collaborative editing of the resources by many users concurrently.

5.2 The Interface

Our intent in this work has been to create a user-friendly interface that would facilitate users in the building of diachronic lexica using the *lemonDia model* without that is involving them too deeply in the details of the formal model underlying the representation. Our interface is web-based and supports the creation both of linked data lexica and related temporal timelines. It accepts as input files in CSV or Excel formats. The rows in these files should include information on (but without being necessarily limited to) written forms, lemmas, roots, etymologies, collocations, meanings, semantic shifts, and the time intervals in which each sense was used. Before importing this input file, it is necessary to enter information about the different time periods listed in the CSV, e.g., OE1. The interface allows users to specify how many time intervals to create, and each interval can be specified by means of the pop-up panel shown in Figure 3.

A lexicon is then generated using the data contained in the input file. Afterwards it is possible to export this lexicon in various formats, such as RDF/XML and TURTLE, or JSON, in order to use it, for example, as input to one of the stages in a Natural Language Processing pipeline. Once the lexicon has been created, the user can visualize all semantic shifts using a graphical mode that uses a timeline graph. Figure 2 shows an example. The interface is composed of three panels. The “lexicon info” panel shows the time intervals that make up the period covered by the evolution of the language; the “model” area contains the lexicon; and the “semantic shift” area shows the temporal evolution of the senses in the lexicon. Figure 2 gives the timelines of seven senses belonging to the lexical field of the word “shame”: so that we can see, for example, the usage of the sense “to become red” is attested between 950 and 1050 in the old english corpus, and the semantic shift type is a metonomy.

From the technical point of view, the tool is based on a software design pattern known as “three-tier architecture”, and exploits Apache Tomcat v7.0 as a web server. The system was implemented using the Java 2 Standard Edition (J2SE) framework which allows the easy manipulation of unicode characters and can be extended to other languages using different writing systems. The OWLAPI has been used for the management of the *lemonDia* model. The presentation tier has been implemented by means of Java Server Faces (JSF) and Primefaces v5.1. This technology allows concurrent access to the imported lexicon and in subsequent versions, we are planning to add functionality that will allow more than one user to carry out management and query tasks on the lexicon at the same time.

6 Future Work

In this article we have shown how the *lemonDia* model can facilitate the creation and subsequent querying of temporal information in diachronic lexical linked open datasets. Furthermore, we have described a preliminary version of a user-friendly interface that assists non expert users in the creation of diachronic lexica. Our tool allows users to import a CSV or Excel file containing lexical data and to subsequently encode the lexicon in RDF using the *lemonDia* model, as well as browsing the temporal information associated with word senses in the lexicon.

We are planning on developing a first release of this tool as an open source application in the near future. In subsequent work we would like to concentrate on the following four aspects: i) extending our tool in order to support the management and editing of the imported lexicon; ii) enhancing the tool with query capabilities by means of a controlled natural language query interface; iii) enabling the importation of ontologies and the association of ontological concepts with individual word senses; iv) extending the *lemonDia* model with the attestations of the word in the corpus.

Acknowledgements

We would like to thank Javier Díaz-Vera for making the Old English emotion dataset available to us.

References

James F Allen. 1983. Maintaining knowledge about temporal intervals. *Communications of the ACM*, 26(11):832–843.

Sotiris Batsakis, Euripides Petrakisb, Ilias Tachmazidisa, and Grigoris Antonioua. 2009. Temporal representation and reasoning in owl 2.0. *Semantic Web Journal.*

Nuria Bel, Sergio Espeja, Montserrat Marimon, and Marta Villegas. 2008. Coldic, a lexicographic platform for lmf compliant lexica. In Bente Maegaard Joseph Mariani Jan Odijk Stelios Piperidis Daniel Tapias Nicoletta Calzolari (Conference Chair), Khalid Choukri, editor, *Proceedings of the Sixth International Conference on Language Resources and Evaluation (LREC'08)*, Marrakech, Morocco, may. European Language Resources Association (ELRA). http://www.lrec-conf.org/proceedings/lrec2008/.

Paul Buitelaar, Philipp Cimiano, Peter Haase, and Michael Sintek. 2009. Towards linguistically grounded ontologies. In *Proceedings of the 6th European Semantic Web Conference on The Semantic Web: Research and Applications*, ESWC 2009 Heraklion, pages 111–125, Berlin, Heidelberg. Springer-Verlag.

Philipp Cimiano, Paul Buitelaar, John McCrae, and Michael Sintek. 2011. Lexinfo: A declarative model for the lexicon-ontology interface. *Web Semantics: Science, Services and Agents on the World Wide Web*, 9(1).

Javier E Díaz-Vera. 2014. From cognitive linguistics to historical sociolinguistics: The evolution of old english expressions of shame and guilt. *Cognitive Linguistic Studies*, 1(1):55–83.

Judith Eckle-Kohler, John McCrae, and Christian Chiarcos. 2015. lemonUby - A large, interlinked, syntactically-rich lexical resource for ontologies. *SEMANTIC WEB*, 6(4):371–378.

Gil Francopoulo, Monte George, Nicoletta Calzolari, Monica Monachini, Nuria Bel, Y Pet, and Claudia Soria. 2006. Lexical markup framework (lmf. In *In Proceedings of LREC2006*.

Aldo Gangemi. 2005. Ontology design patterns for semantic web content. In *Proceedings of the 4th International Conference on The Semantic Web*, ISWC'05, pages 262–276, Berlin, Heidelberg. Springer-Verlag.

Thomas M. Johnson, Harold R. Solbrig, Daniel C. Armbrust, and Christopher G. Chute. 2005. Lexgrid editor: Terminology authoring for the lexical grid. In *AMIA 2005, American Medical Informatics Association Annual Symposium, Washington, DC, USA, October 22-26, 2005.*

Tom Kenter, Tomaž Erjavec, Maja žorga Dulmin, and Darja Fišer. 2012. Lexicon construction and corpus annotation of historical language with the cobalt editor. In *Proceedings of the 6th Workshop on Language Technology for Cultural Heritage, Social Sciences, and Humanities*, LaTeCH '12, pages 1–6, Stroudsburg, PA, USA. Association for Computational Linguistics.

Fahad Khan, Federico Boschetti, and Francesca Frontini. 2014. Using lemon to Model Lexical Semantic Shift in Diachronic Lexical Resources. Proceedings of the Workshop on Linked Data in Linguistics 2014 (LDL-2014).

Fahad Khan, Javier E. Díaz-Vera, and Monica Monachini. 2016. Representing polysemy and diachronic lexico-semantic data on the semantic web ? In Isabelle Draelants, Catherine Faron-Zucker, Alexandre Monnin, and Arnaud Zucker, editors, *Proceedings of the Second International Workshop on Semantic Web for Scientific Heritage co-located with 13th Extended Semantic Web Conference (ESWC 2016), Heraklion, Greece, May 30th, 2016.*, volume 1595 of *CEUR Workshop Proceedings*, pages 37–46. CEUR-WS.org.

John P. McCrae and Christina Unger. 2014. Design patterns for engineering the ontology-lexicon interface. In Paul Buitelaar and Philipp Cimiano, editors, *Towards the Multilingual Semantic Web*, pages 15–30. Springer.

John McCrae, Guadalupe Aguado-de Cea, Paul Buitelaar, Philipp Cimiano, Thierry Declerck, Asunción Gómez Pérez, Jorge Gracia, Laura Hollink, Elena Montiel-Ponsoda, Dennis Spohr, et al. 2010. The lemon cookbook.

John McCrae, Elena Montiel-Ponsoda, and Philipp Cimiano, 2012. *Integrating WordNet and Wiktionary with lemon*, pages 25–34. Springer.

Elena Montiel-Ponsoda, Guadalupe Aguado-de Cea, Asuncin Gmez-Prez, and Wim Peters. 2008. Modelling multilinguality in ontologies. Poster, CoLing 2008.

Elena Montiel-Ponsoda, J. McCrae, and Philipp Cimiano. 2012. Collaborative semantic editing of linked data lexica. In *Proceedings of the Eight International Conference on Language Resources and Evaluation (LREC'12)*, pages 2619–2625. European Language Resources Association (ELRA).

Martin O'Connor and Amar Das. 2009. Sqwrl: A query language for owl. In *Proceedings of the 6th International Conference on OWL: Experiences and Directions - Volume 529*, OWLED'09, pages 208–215, Aachen, Germany, Germany. CEUR-WS.org.

Jacquelijn Ringersma and Marc Kemps-Snijders. 2007. Creating multimedia dictionaries of endangered languages using LEXUS. In *INTERSPEECH 2007, 8th Annual Conference of the International Speech Communication Association, Antwerp, Belgium, August 27-31, 2007*, pages 1529–1532.

Rim Touhami, Patrice Buche, Juliette Dibie-Barthélemy, and Liliana Ibănescu, 2011. *An Ontological and Terminological Resource for n-ary Relation Annotation in Web Data Tables*, pages 662–679. Springer Berlin Heidelberg, Berlin, Heidelberg.