Towards a Module for Lexicography in OntoLex

Julia Bosque-Gil, Jorge Gracia, Elena Montiel-Ponsoda

Ontology Engineering Group, Universidad Politécnica de Madrid
jbosque,jgracia,emontiel@fi.upm.es

Abstract. Dictionaries are increasingly being transformed into linguistic linked lata (LLD) relying on the *lemon* and OntoLex models, but this conversion is not always straightforward. For both linked data (LD) based applications to exploit all content provided in dictionaries and lexicographers adopting LD technologies, the original data and structure should be retrievable from the LLD version to prevent any loss of information in the transformation. In this position statement we motivate the need for a new module in OntoLex targeted at the representation of dictionaries and which will address structures and annotations commonly found in lexicography. Some of the issues we identified in our initial experiences are presented as input for discussion, along with our initial approaches to solve them. Such a module is intended to be compatible with other modules in OntoLex and should guarantee information preservation, making LD a viable mechanism for lexicographers in the development of lexica.

Keywords: OntoLex, Linguistic Linked Data, lemon, lexicography, dictionaries

1 Introduction

Over the past few years, more and more efforts are being devoted towards the conversion of dictionaries into Linguistic Linked Data (LLD), based on *lemon* [1] and its more recent version OntoLex[1], a de facto standard to represent ontology-lexica on the Web.These works aim both to enrich the so-called Linguistic Linked (Open) Data cloud[2] with lexical information to be consumed by natural language processing (NLP) tools, and to build bridges between the lexicography and the Semantic Web communities. Recent projects such as LIDER[3], or on-going ones such as ENeL[4], LDH4HELTA[5], and LiODi[6], promote the adoption of linked data technologies in the work with lexicographic resources focusing on language technologies, e-lexicography and linguistic research, respectively. The benefits of

[1] http://www.w3.org/2016/05/ontolex/
[2] http://linguistic-lod.org/llod-cloud
[3] http://www.lider-project.eu/
[4] http://www.elexicography.eu/
[5] http://ldl4.com/
[6] http://acoli.informatik.uni-frankfurt.de/liodi/

representing lexical content from dictionaries as LLD (e.g. interoperability across resources or better visibility and reuse), have already been extensively reported in the literature [2, 3]. By way of illustration, we will refer to the experiences and advantages of migrating monolingual [3], bilingual [4], multilingual [5], Ancient Greek [6], dialectal [7–9], and etymological dictionaries [10, 11], along with the WordNet family of resources [12] and theory-based dictionaries (Pattern Dictionary of English Verbs [13], Parole-Simple Lexica [14, 15]), among others. The added value of using linked data technologies in lexicography and its implications for the micro- and macro structure have been explored as well [16].

Nonetheless, the conversion of a dictionary to OntoLex is not always straightforward. *lemon* was initially developed to enrich a given ontology with a lexical layer, and not with the idea of rendering any already existent dictionary to LLD. A majority of scholars working on this field, however, are turning to *lemon* or OntoLex in pursuit of the latter objective. The more numerous and resource-specific the annotations in a dictionary are, the more complex the modeling solutions are, especially if until then the dictionary was targeted at human users. We are aware that some solutions exceed the needs of lexical information that some NLP tools require. However, if we are also aiming to bring linked data to lexicography, all dictionary content must be taken into account and must be retrievable once converted to LLD, i.e, migrating to LLD should imply no information loss. This means that structural aspects of the dictionary, as for instance senses and homographs order, along with the sub-sense hierarchy some dictionaries display, should be kept in mind when offering modeling solutions. There is a range of dictionary annotations (domain of usage, region, frequent use tags, restrictions on number and gender depending on a sense, etc.) that affect word meaning and language usage and are not structural in nature. Collocations, idioms, paradigmatic relations, context indicators, semantic selection, among other aspects, are presented differently in dictionaries and modeling them is not trivial.

The natural doubt that would be entertained by many experts is whether OntoLex is supposed to provide the means to model all aspects of a dictionary or whether this is outside of its scope, ontology lexicalization, and therefore should be tackled by another initiative. In this paper we motivate our insights on OntoLex to enable dictionary representation as LLD in all its granularity, and advocate for the creation of a lexicography-specific module that would gather elements concerning dictionary structure and annotations. The module could also link to other potential modules that might be proposed, such as an etymology-oriented one to support etymological dictionaries, for instance.

The rest of the paper is organized as follows: Section 2 goes through the state of the art on LLD and lexicography and some of the problems encountered during the representation of dictionaries as LLD. Our motivation for OntoLex to be able to tackle those and the issues presented throughout the paper is stated in that section as well. Section 3 describes five of a series of issues we identified in our work modeling and analyzing dictionary entries, and which we argue serve as input for discussion on the need for a module for lexicography. Our initial approaches towards such a module and a description of how it would solve the

described issues are outlined in Section 4, while Section 5 offers some concluding remarks.

2 Background and Motivation

There have been several reports in the literature on the conversion of dictionaries to LLD, most of them relying on *lemon* or OntoLex. However, proprietary formats, such as that of K Dictionaries (KD)[7], often have XML tags used in their annotation schemes that refer to linguistic categories or features which are not present in available repositories of linguistic categories or which lack a compatible definition that prevents us from reusing the ontology entity at hand. Ad hoc vocabularies were defined to migrate content from the German monolingual dictionary of KD's Global Series [3] and its Spanish multilingual set [5]. These works approached issues which affect, for example, the relation between a lexical sense and the lexicalized phrases and idioms in which it occurs, regional restrictions, lexical and semantic selection (in general) of lexical entries, groups of homographs, tone and register indications, inflection groups, context of use, frequency modifiers to register, etc. Multilingual dictionaries pose further problems due to the modeling of examples and translations of examples, as well as alternative forms of those translations (e.g. an example in English translated to Japanese in kanji and hiragana, and that translation in turn with a transliteration in rōmaji). The set of thirteen dictionaries (dialectal, bilingual, monolingual, historical, etc.) converted as part of the ENel Action [2] required the definition of new properties to encode different types of temporal information and etymological aspects.

Structures typically found in dictionaries, such as the sense and sub-sense hierarchy in an entry, are not trivial to model either. *polyLemon* [6], developed as part of the conversion of the Liddell-Scott Greek-English Lexicon to *lemon*, was suggested in order to capture the sense and sub-sense structure in dictionaries using properties such as `senseChild` and `senseSibling` to relate senses and their parent senses in the dictionary entry.

The accurate representation of etymological information as LLD is key in the conversion of historical and etymological dictionaries. An extension to *lemon*, *lemonet*, to represent etymological information of lexical entries was proposed [10] and, more recently, a revisited version builds upon the properties suggested for the modeling of the etymological WordNet[8] to undertake the conversion of the Tower of Babel (Starling) in the LiODi project [11]. Some recent work on the conversion of the classical Arabic Dictionary Al-Qamus to *lemon* and LMF has been undertaken [17], but no pointers or traceback to the original structure are given in the work.

Alternatives to the use of OntoLex are available as well. The Oxford Global Languages Ontology (OGL) [18] has been developed to model and integrate multilingual linguistic data from Oxford Dictionaries and emerges as an ontology

[7] `http://kdictionaries-online.com/`
[8] `http://www1.icsi.berkeley.edu/~demelo/etymwn/`

exclusively created to meet dictionary representation requirements. It accounts for a range of information found in dictionaries, from inflected forms to semantic relations, pragmatic features and etymological data. The focus is laid on the representation of grammatical information with cross-linguistic validity and the respect towards grammar traditions. However, some modeling decisions and class definitions differ from those suggested in the OntoLex core (e.g. `Form` in OntoLex vs. a `Form` in OGL) and the emphasis is not set on the reuse of available ontology entities.

In this position paper we do not focus on a particular kind of lexical information present in dictionaries (e.g., etymology or morphology) but we aim to highlight some difficulties in the modeling of dictionary entries without information loss. Thus, we will not target the representation of resource-specific features of particular dictionaries. Taken into account the problems reported in the literature, and after analyzing dictionary entries in e-dictionaries of English (Oxford [19], Merriam Webster [20], American Heritage Dictionary [21], COBUILD Advanced English Dictionary and Collins English Dictionary [22]), German (Duden [23], PONS Deutsch als Fremdsprache [24]), and Spanish (Diccionario de la Lengua Española (DLE) [25], CLAVE [26]), we report on some of the issues we gathered which may pose problems for the modeling with OntoLex and which we believe call for the definition of a new module to account for them. Future steps include the analysis of dictionaries in languages that are underrepresented in the LLOD cloud (e.g. Japanese) to identify further representation challenges.

We ground our proposal for a lexicography module on the following four points: (1) the use of OntoLex by the majority of the community to convert linguistic resources to LLD instead of to lexicalize ontologies, (2) the nature of *lemon* being descriptive but not prescriptive and the respect towards different lexicographic views, (3) the coming together of the lexicography and the Semantic Web communities and potential benefits that LLD may bring about to lexicography, assuming it involves no information loss, and (4) the reuse of already available mechanisms in OntoLex.

3 Issues

In the following we report on some of the issues we have come across after our experiences in converting dictionaries to LLD and our analysis of dictionary entries in English, German and Spanish. Here we restrain ourselves to issues that reveal current limitations of the OntoLex model, i.e, cases in which applying the *lemon* core implies a different view on the data than the one provided in the original resource and, therefore, an information loss (type 1, hence T1), and missing entities, e.g. a property or a class, to account for information mostly found in dictionaries (type 2, hence T2). We have already raised some of these issues as input for discussion to the OntoLex community.[9]

[9] https://www.w3.org/community/ontolex/wiki/Lexicography

Issue 1 (T1). Headwords that can take different parts-of-speech
Both *lemon* and OntoLex specify a lexical entry as a word, a multiword expression or an affix with a single part-of-speech, morphological pattern, etymology and set of senses.[10] However, a headword in a dictionary may occur with different parts-of-speech depending on context and its senses are nonetheless defined in the same dictionary entry, all of them derived from the same etymology (no homonymy involved). Applying the OntoLex model would imply the generation of several `ontolex:LexicalEntr[ies]`, one per each part-of-speech the headword can take. Splitting the dictionary entry into several lexical entries would cause loss of information (shared etymology, pronunciation, senses implicitly related) and does not keep track of the dictionary representation. Examples: *poison, bread, water* (noun and verb), Sp. *lento* 'slow, slowly' (adjective and adverb), Sp. *alto* 'tall, loudly, height' (adjective, adverb and noun).

Issue 2 (T1). Lexical sense requiring a particular form
Some senses of a dictionary headword require a particular form, e.g. in English a plural form or in Spanish a masculine or feminine one. Since the meaning in these cases is associated with the form and it may differ significantly from other senses that do share gender or number features, splitting the dictionary entry into different lexical entries would be an option (see Issue 1). An alternative is the linking of that sense to elements in a catalog of grammatical categories which encode those grammatical restrictions, but we would need an exhaustive list of them in order for this option to be applicable. Examples: *refreshment(s)*, Sp. *cometa* (m.) 'comet', (f.) 'kite'. In these cases, the dictionary entry can be a single one (e.g. *refreshment* in English or *cometa* in Spanish) but one of its senses indicates a preferred form. In the case of *refreshment*, the plural form is used if the intended meaning is snacks and beverages; with the Spanish *cometa*, the feminine form is applicable when referring to a kite, the masculine when denoting a comet. Further examples are *good(s)*, *manner(s)*; Sp. *frente* (m.) 'front', (f.) 'forehead'.

Issue 3 (T2). Usage examples and their translations
Usage examples of a word or multiword expression are often provided in the definition of each of a dictionary entry's senses. Lexinfo includes a property `lexinfo:senseExample` to describe an example of a sense (as a subproperty of `lemon:definition`) and which is linked to the example data category in ISOCat.[11] Nonetheless, due to it being a datatype property, it does not enable including further information on the example or to establish translation relations among examples, which is common practice in bilingual and multilingual dictionaries. The *lemon* model included a `lemon:UsageExample` class and a property `lemon:example` to link to it, but OntoLex does not cover this aspect yet. Exam-

[10] `https://www.w3.org/2016/05/ontolex/#lexical-entries`
[11] `http://www.isocat.org/`

ples: Sp. *Preocuparse* 'worry'; Sp. *no hay por qué preocuparse* 'there is nothing to worry about' (Collins English-Spanish Dictionary).[12]

Issue 4 (T2). Sense and homographs order
The order of senses may be based on frequency of use, date of origin, concreteness (from the most concrete to most abstract sense, etc.). Homographs are also given according to some ordering criteria that may vary from dictionary to dictionary. Their order should be searchable and retrievable as to recover the information provided in the original resource. Examples: *Boa*: noun. (1) any of a family (Boidae) of large snakes that kill by constriction and that includes the boa constrictor, anaconda, and python (2) a long fluffy scarf (Merriam Webster Dictionary)[13]; *bat*[1]: n. 1. A stout wooden stick; a cudgel [...]; *bat*[2]: n. Any of various nocturnal flying mammals of the order Chiroptera [...] (American Heritage Dictionary).[14]

Issue 5 (T2). Semantic selection
Some dictionaries indicate the semantic features of the lexical items that an entry (in one of its senses) selects or even the exact lexical items with which it collocates. This is usually indicated either with a specific tag (e.g. KD's tag *Range Of Application*), or in-between parentheses at the beginning of a definition. Examples are, for instance, the dictionary entry for the German verb *dämmen*, which in its sense 'to insulate, absorb, mute' selects arguments that denote warmth or sound (German *Wärme, Schall*, etc.) (KD)[15], the adjective *cozy*, meaning *beneficial to all those involved and possibly somewhat corrupt* if predicated from a transaction or an arrangement (Google Dictionary)[16]; or the collocational measure words of *luck*: *stroke, piece of* (Oxford Collocations Dictionary).[17] The `synsem:OntoMap` class allows to map a syntactic frame to an ontology entity, so that the frame and its arguments are linked to the ontology elements that they lexicalize. Even though dictionaries commonly include information on subcategorization (transitive/intransitive/reflexive etc. annotations for verbs, for instance), details on the syntactic frame are not always provided beyond those annotations. Since in dictionary conversion we often lack a given ontology and detailed syntactic information is not provided, the mapping between syntactic arguments and ontology entities seems difficult to establish automatically via

[12] "worry". Intransitive Verb. Def. 1, Example 5. Collins English-Spanish Dictionary. `http://www.collinsdictionary.com/dictionary/english-spanish/worry`. Last accessed 13/07/2017.

[13] "boa". n. Merriam Webster Dictionary. `http://www.merriam-webster.com/dictionary/boa`. Last accessed 13/07/2017. Example of logical order of senses inspired by Diccionario de la Lengua Española, Guía de Consulta, `http://dle.rae.es/`.

[14] "bat"[1]. n. Def. 1.; "bat"[2]. n. Def. American Heritage Dictionary, `https://ahdictionary.com/word/search.html?q=bat`. Last Acessed 13/07/2017.

[15] "dämmen". Sense 2. K Dictionaries Global Series Multilingual German Dictionary

[16] "cosy". Def. 1., sub-sense 3. Google Dictionary. Last Accessed 13/07/2017.

[17] "luck". QUANT. Deuter, M. (Ed.). (2008). Oxford Collocations Dictionary for students of English. Oxford University Press, p. 473.

`synsem:OntoMap`: how do we automatically represent that the adjective *cozy* has a meaning only applied to *transaction* or *agreement* or that the measure words that collocate with *luck* are *stroke, piece* if the morphosyntactic information provided in the dictionary is just that *cozy* is an adjective and *luck* a noun? Furthermore, `synsem:condition` (in its turn subsuming `synsem:propertyRange` and `synsem:propertyDomain`) enables us to state constraints on the arguments of a predicate in a given ontology.[18] The possibility of reusing it to state the constraints on syntactic arguments even in cases in which we lack a given ontology and therefore are not mapping to given ontology properties has to be further analyzed. In addition, the potential links between the modeled entries (e.g. *piece* and *luck*) (i.e, the links at the lexical level) are also to be considered, for instance, by taking into account recent proposals on the representation of lexical functions as LLD [27].

4 A Module for Lexicography

The previous section dealt with some of the issues we encountered in our work with dictionaries and the potential ones that may rise with other lexicographic works that have not been migrated to LLD yet. In the following we draft a potential solution which can serve as basis for a new module in OntoLex specifically developed for the representation of dictionaries after thorough revision and improvement according to the community's feedback.

In order to keep track of the dictionary representation and prevent any loss of information mentioned in Issue 1, related to the splitting of dictionary entries in several lexical entries, we propose a new class `DictionaryEntry`. This new class would both enable to group together lexical entries as well as to associate any information shared by all of them. A class `Entry` was proposed in the recent Oxford Global Languages Ontology [18] (OGL) to store provenance data and allow a fast filtering, whereby a lexical entry in OGL would be linked to all OGL entries that provide information about it. In our view, we distinguish lexical entries and lexicons (as containers of lexical entries), from the original dictionary entry (a new class `DictionaryEntry`) and the original dictionary resource (`Dictionary`), which would serve in turn to record the provenance of each dictionary entry. Mirroring the `lime:Lexicon-LexicalEntry` relation we suggest a `Dictionary-DictionaryEntry` one. Any lexical entry created during the conversion to LLD but not originally provided in the resource would then belong to a `lime:Lexicon`, but not to the instance of `Dictionary` representing that resource. A `lime:Lexicon` in English, for example, could aggregate lexical entries created on the fly by the LLD expert or original ones coming from as many English dictionaries as desired. These dictionaries can in turn differ in their modeling and their views on the data, their criteria of sense ordering or their structure.

Regarding Issue 2, the `DictionaryEntry` class would allow to divide a single lexical entry into several ones if desired, each with a different preferred form,

[18] `https://www.w3.org/2016/05/ontolex/#conditions`

while maintaining the original dictionary representation. If the dictionary entry is not split, the option of linking a sense to a grammatical restriction on gender or number from an external catalog would solve the issue, although the implications of this solution (its benefits and drawbacks) will need further analysis.

In order to represent usage examples and their translations (Issue 3) we propose to go back to `lemon:UsageExample` and link it to a LexicalSense. A new class `ExampleCluster` would link to `UsageExamples` that are translations from each other. The use of the `vartrans` module to model translations among senses would imply the creation of lexical senses for each example, and therefore treating the example as a lexical entry, which we deem is beyond the definition of lexical entries.

Issue 4 was concerned with the order of senses in a dictionary entry and the order of homographs in the macrostructure of the dictionary. There are different possible approaches to resolve this: reusing already available RDF mechanisms, reifying the sense order in a new class `SenseOrder`, or defining a new property `senseOrder` attached to the lexical sense. The first and straightest forward option involves the reuse of `rdfs:Container`s to declare with e.g. `rdf:_1, rdf:_2` that a particular sense is the first or the second one. However, cases in which a set of senses allows for various orderings, depending on the ordering criterion, or in which some senses come from different dictionaries (each one with its order), should also be accounted for.

The second option suggests that the sense order is reified in a class `SenseOrder` linked to the lexical sense. This class would enable us to record the position of that sense, its provenance (presumably an instance of the class `Dictionary`), and, if desired, the ordering criterion. If repeated senses were identified (e.g. senses that share a definition in both dictionaries), `SenseOrder` would allow us to have one single lexical sense with two different positions according to the two different orderings and dictionaries, in a similar fashion as two containers with two different sequences of senses. Alternatively, if we assume that a lexical sense always comes from just one dictionary source, a property `senseOrder` would suffice.

Issue 5, dealing with semantic selection, has been brought up for further discussion in this paper as to see whether it could be covered by *synsem* module mechanisms or whether it would require new entities in the context of the lexicography module. As part of the conversion of the KD's Global Series Spanish Multilingual Dictionary [5], the semantic selection information provided by KD's tag *Range of Application* was captured by the use of `synsem:condition`. In that approach, `synsem:condition` would link a lexical sense to a blank node[19] with an `rdf:value` recording the strings given as arguments in the original data. This modeling allowed us to deal with the lack of a given ontology and detailed information on the syntactic frames of lexical entries for each of their senses. Thus, the focus was set on representing the data just as it was in its original format while being compliant with the OntoLex formal specification and reusing its elements as much as possible. We argue that the lexicography module should

[19] `synsem:condition` has `rdfs:Resource` as range.

aim to set the basis to exploit at the dictionary's macro-structure level the potential benefits of establishing semantic relations among lexical senses based on lexical selection or among syntactic frames and arguments and the ontology entities that they denote. To this aim, overcoming the lack of detailed syntactic information in the dictionary as well as the lack of a given ontology to lexicalize becomes essential.

5 Conclusion

OntoLex is increasingly being used to convert linguistic resources to LLD outside the scope of ontology localization. In this position statement we have drawn attention to a series of issues raised in the literature on LLD related to the conversion of dictionaries to LD and to five of the ones we came across in the same line of work and after a later analysis of several additional dictionaries. We argue that the OntoLex model should enable the preservation of the content and the structure of the original resource, even if the LLD expert opts for a different representation more suited to the exploitation of the data by external applications or more in line with his or her view on the lexicon-ontology interface. We have outlined some of our insights on how to address these issues in a new module for lexicography. It would be compatible with the mechanisms suggested in the state-of-the art on dictionaries in LLD, as of the moment of writing, and also with other potential modules for the representation of specific lexical aspects (e.g. etymology). The final module is intended to be dictionary-agnostic in the sense that it should be applicable (and combined with other modules if necessary) to different kinds of dictionaries (e.g., general, collocations, learner's, etymological, historical, etc.). This would bring linked data (LD) closer to lexicography not only with the aim of leveraging already available dictionaries in LD for NLP tasks, but also of introducing linked data in the work carried out in that discipline.

Acknowledgments. We thank the anonymous reviewers for their insightful suggestions that have contributed to the improvement of this article. This work is supported by the Spanish Ministry of Education, Culture and Sports through the Formación del Profesorado Universitario (FPU) program, and by the Spanish Ministry of Economy and Competitiveness through the project 4V (TIN2013-46238-C4-2-R) with the FEDER funding scheme, the Juan de la Cierva program, and the Excellence Network ReTeLe (TIN2015-68955-REDT).

References

1. McCrae, J., Aguado-de-Cea, G., Buitelaar, P., Cimiano, P., Declerck, T., Gómez-Pérez, A., Wunner, T.: Interchanging lexical resources on the Semantic Web. Language Resources and Evaluation, 46(4), 701719 (2012)

2. Declerck, T., Wandl-Vogt, E., and Mörth, K.: Towards a Pan European Lexicography by Means of Linked (Open) Data. In: Electronic lexicography in the 21st century: linking lexical data in the digital age. Proceedings of the eLex 2015 conference, pp. 342–355. (2015)
3. Klimek, B., and Brümmer, M.: Enhancing lexicography with semantic language databases. Kernerman Dictionary News 23, 5–10 (2015)
4. Gracia, J., Villegas, M., Gómez-Pérez, A., and Bel, N.: The Apertium Bilingual Dictionaries on the Web of Data. Semantic Web Journal, IOS Press (2016)
5. Bosque-Gil, J., Gracia, J., Montiel-Ponsoda, E., and Aguado-de-Cea, G.: Modelling multilingual lexicographic resources for the Web of Data: The K Dictionaries case. In: GLOBALEX 2016 Lexicographic Resources for Human Language Technology Workshop Programme, pp. 65. (2016)
6. Khan, F., Díaz-Vera, J. E., and Monachini, M.: Representing Polysemy and Diachronic Lexicosemantic Data on the Semantic Web. In: Proceedings of the Second International Workshop on Semantic Web for Scientific Heritage co-located with 13th Extended Semantic Web Conference (ESWC 2016), Heraklion, Greece, pp. 37–46. (2016)
7. Declerck, T., and Mörth, K.: Towards a Sense-based Access to Related Online Lexical Resources. In: Proceedings of Euralex 2016, pp. 342–355. (2016)
8. Declerck, T., and Wandl-Vogt, E.: How to semantically relate dialectal Dictionaries in the Linked Data Framework. In: EACL 2014, p. 9 (2014)
9. Declerck, T., and Wandl-Vogt, E.: Cross-linking Austrian dialectal Dictionaries through formalized Meanings. In: Proceedings of the XVI EURALEX International Congress, pp. 329–343 (2014)
10. Chiarcos, C., and Sukhareva, M.: Linking Etymological Databases. A Case Study in Germanic. In: 3rd Workshop on Linked Data in Linguistics: Multilingual Knowledge Resources and Natural Language Processing, p. 41. (2014)
11. Abromeit, F., and Fäth, C.: Linking the Tower of Babel: Modelling a Massive Set of Etymological Dictionaries as RDF. In: LDL 2016 5th Workshop on Linked Data in Linguistics: Managing, Building and Using Linked Language Resources, p. 11 (2016)
12. McCrae, J., Fellbaum, C., and Cimiano, P.: Publishing and Linking WordNet using lemon and RDF. In: Proceedings of the 3rd Workshop on Linked Data in Linguistics. (2014)
13. El Maarouf, I., Bradbury, J., and Hanks, P.: PDEV-lemon: a Linked Data implementation of the Pattern Dictionary of English Verbs based on the Lemon model. In: Proceedings of the 3rd Workshop on Linked Data in Linguistics (LDL): Multilingual Knowledge Resources and Natural Language Processing at the Ninth International Conference on Language Resources and Evaluation (LREC14), Reykjavik, Iceland. (2014)
14. Del Gratta, R., Frontini, F., Khan, F., and Monachini, M.: Converting the Parole Simple Clips Lexicon into RDF with *lemon*. Semantic Web, 6(4), 387–392 (2015)
15. Villegas, M., and Bel, N.: PAROLE/SIMPLE lemonontology and lexicons. Semantic Web 6(4), 363–369 (2015)
16. Bosque-Gil, J., Gracia, J., and Gómez-Pérez, A.: Linked data in lexicography. Kernerman Dictionary News, 24, 19 – 24 (2016)
17. Khalfi, M., Nahli, O., and Zarghili, A.: Classical dictionary Al-Qamus in lemon. In: 2016 4th IEEE International Colloquium on Information Science and Technology (CiSt), pp. 325–330. (2016)

18. Parvizi, A., Kohl, M., González, M., Saurí, R.: Towards a Linguistic Ontology with an Emphasis on Reasoning and Knowledge Reuse. In: International Conference on Language Resources and Evaluation (LREC16) (2016)
19. Oxford Living Dictionaries, http://en.oxforddictionaries.com/. Last accessed 13/07/2017.
20. Merriam-Webster Dictionary, http://www.merriam-webster.com/. Last accessed 13/07/2017.
21. American Heritage Dictionary, http://ahdictionary.com/. Last accessed 13/07/2017.
22. COBUILD Advanced English Dictionary and Collins English Dictionary, http://www.collinsdictionary.com/. Last accessed 13/07/2017.
23. Duden Online Wörterbuch, http://www.duden.de/woerterbuch. Last accessed 13/07/2017.
24. PONS Online- Wörterbuch Deutsch als Fremdsprache, http://de.pons.com/. Last accessed 13/07/2017.
25. Real Academia Española, Diccionario de la Lengua Española, 23^a Edición, 2014 http://dle.rae.es. Last accessed 13/07/2017.
26. SM diccionarios. Diccionario Clave (online), http://clave.smdiccionarios.com/app.php. Last accessed 13/07/2017.
27. Fonseca, A., Sadat, F., and Lareau, F.: Lexfom: a lexical functions ontology model. COLING 2016, 145 (2016)