

Towards a terminological knowledge base on Islamic archaeology: linguistic and conceptual aspects

Bruno Almeida & Rute Costa

Atelier DAHLIA@ECG2020

Brussels, 28 Jan. 2020

Summary

1. Terminology and archaeology of al-Andalus
2. Overview of OntoAndalus
3. Representation of Portuguese and Spanish artefact terms
4. Conclusions

Archaeology of al-Andalus

- Al-Andalus: territory in the Iberian Peninsula under Islamic rule (711-1492)
- Wide variety of cultural remains, in which pottery sherds are predominant



Importance of terminology work in the domain

*The scientific community has faced **difficulties in nomenclature**, with emphasis on a clear lack of definition and disorder in the attribution of a **unique and consensual name** to each of the artefacts.*

[...]

*A reflection on terminology has shown, in Spain, that this methodology can open **new paths of research**, namely in the field of functionality of objects. On the other hand, the **computerization of data**, in order to be effective, requires rigor and precision of the terms.*

Torres, Gómez Martínez, & Ferreira, 2003, p. 125 (English translation by DeepL)

Terminological knowledge bases

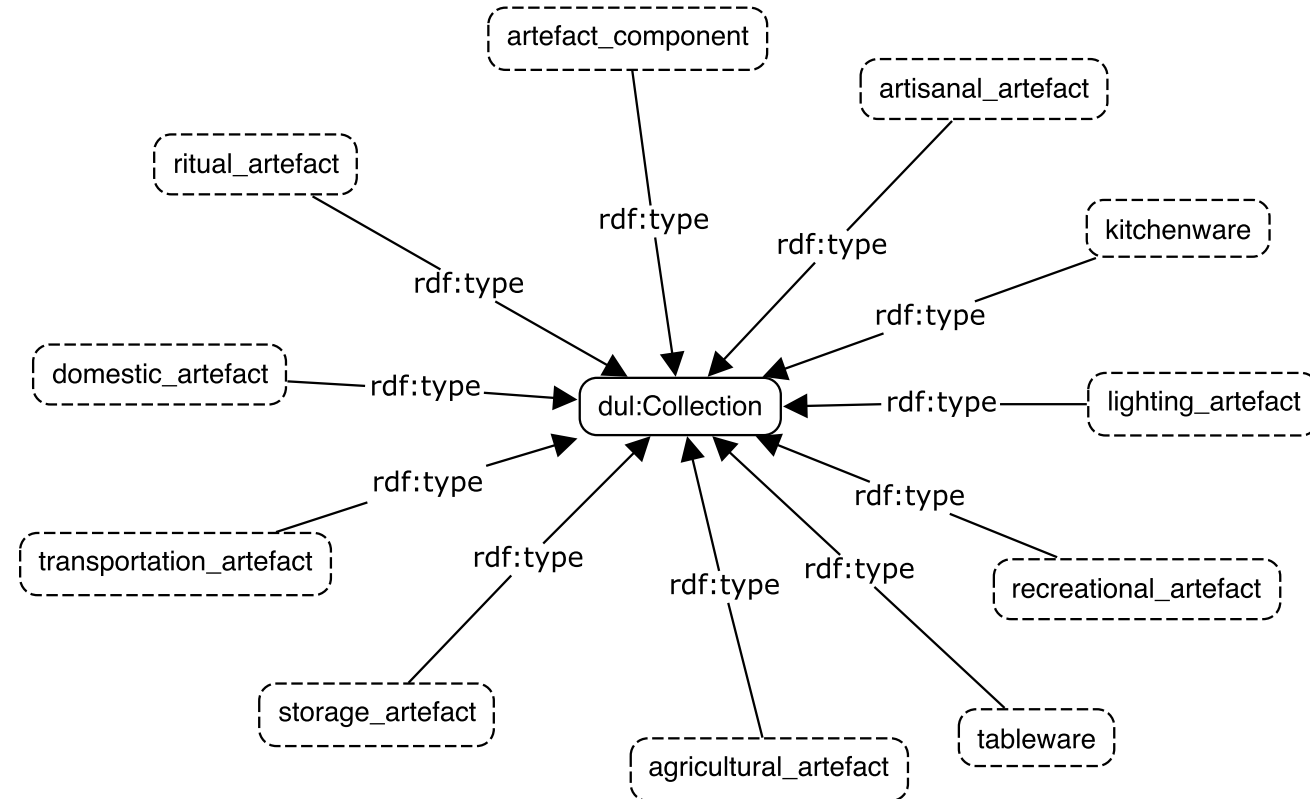
- Resources with structured linguistic and conceptual information about specialised domains
 - Multilingual terminology
 - Concepts and other units of knowledge
 - Natural language definitions based on formal definitions of concepts
- Interdisciplinary work involving terminologists/linguists, engineers and domain specialists
- Closer ties between terminology work and ontology engineering

Overview of OntoAndalus

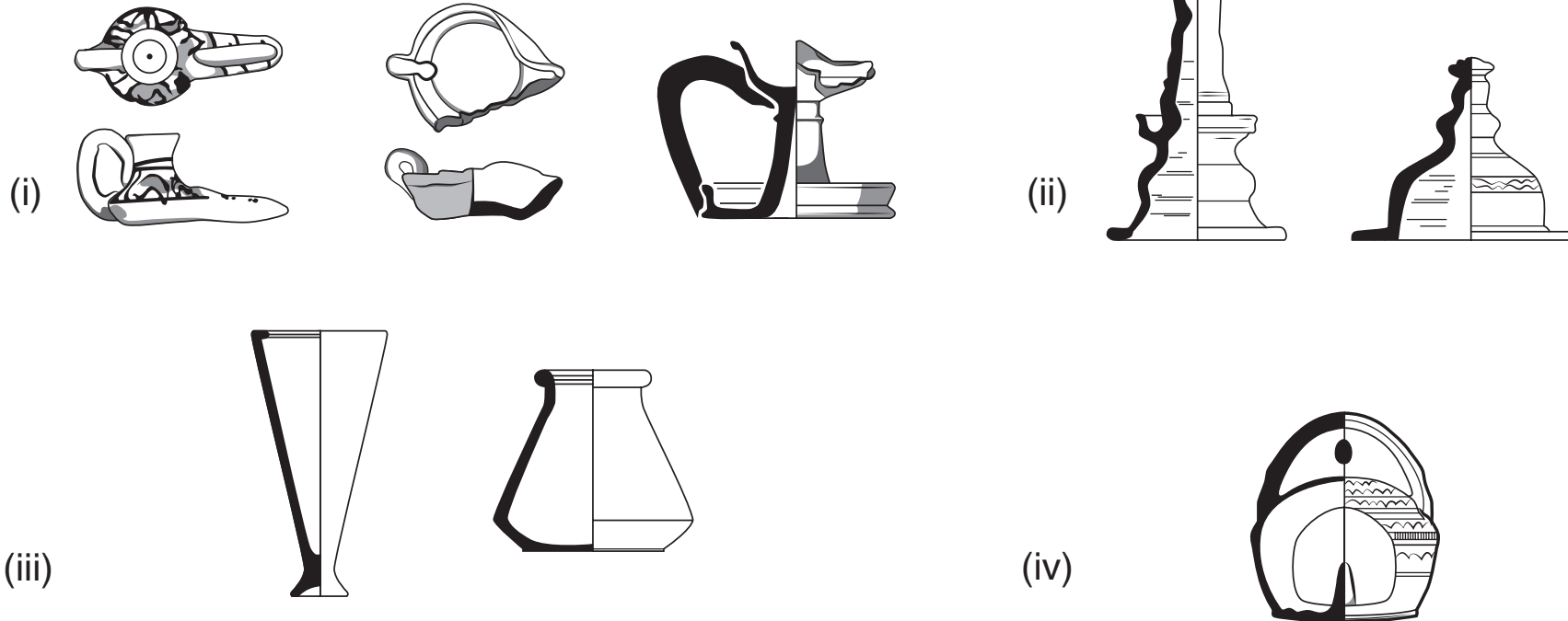
- Ontology of relevant artefacts types in the pottery studies of al-Andalus
- Based on information drawn from a bilingual corpus (Portuguese and Spanish) and English textbooks on archaeology and pottery analysis
- Reuse of DOLCE+DnS Ultralite (DUL) by means of the `owl:imports` property

Typology of artefacts

- 72 artefact types organised in 11 functional collections



Lighting artefacts



Spanish terms: (i) *candil*; (ii) *almenara*, *policandela*; (iii) *lamparilla*; (iv) *fanal*, *linterna*

Describing artefact types

- Definitional pattern: **superordinate class + collection + function + part(s) or component(s)**
- Lamp (*candil*). *Def.* Artefact for lighting in closed spaces composed of at least one spout and a single chamber for liquid fuel.

```
Lamp rdfs:subClassOf dul:DesignedArtifact
```

```
Lamp rdfs:subClassOf dul:isMemberOf some lighting_artefact
```

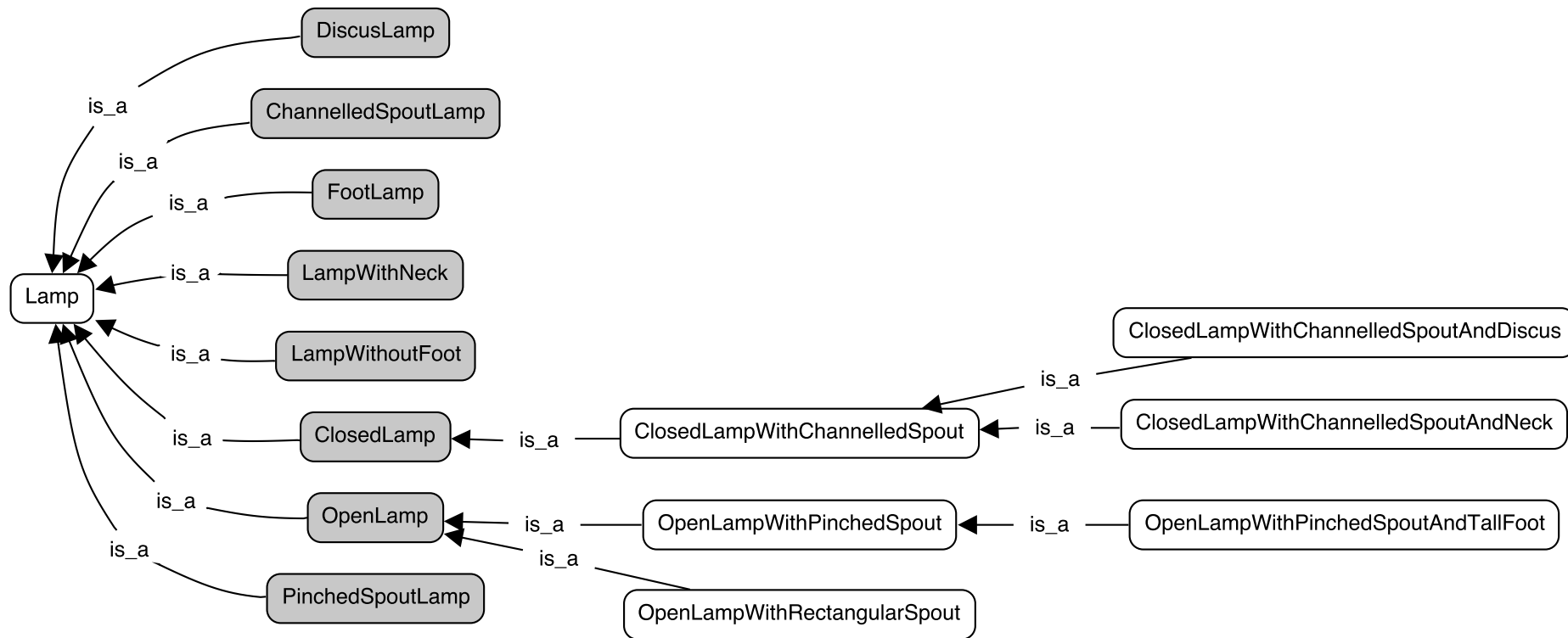
```
Lamp rdfs:subClassOf hasFunction some  
containing_fire_for_lighting_in_closed_spaces_using_liquid_fuel
```

```
Lamp rdfs:subClassOf hasComponent exactly 1 LampFuelChamber
```

```
Lamp rdfs:subClassOf hasComponent min 1 Spout
```

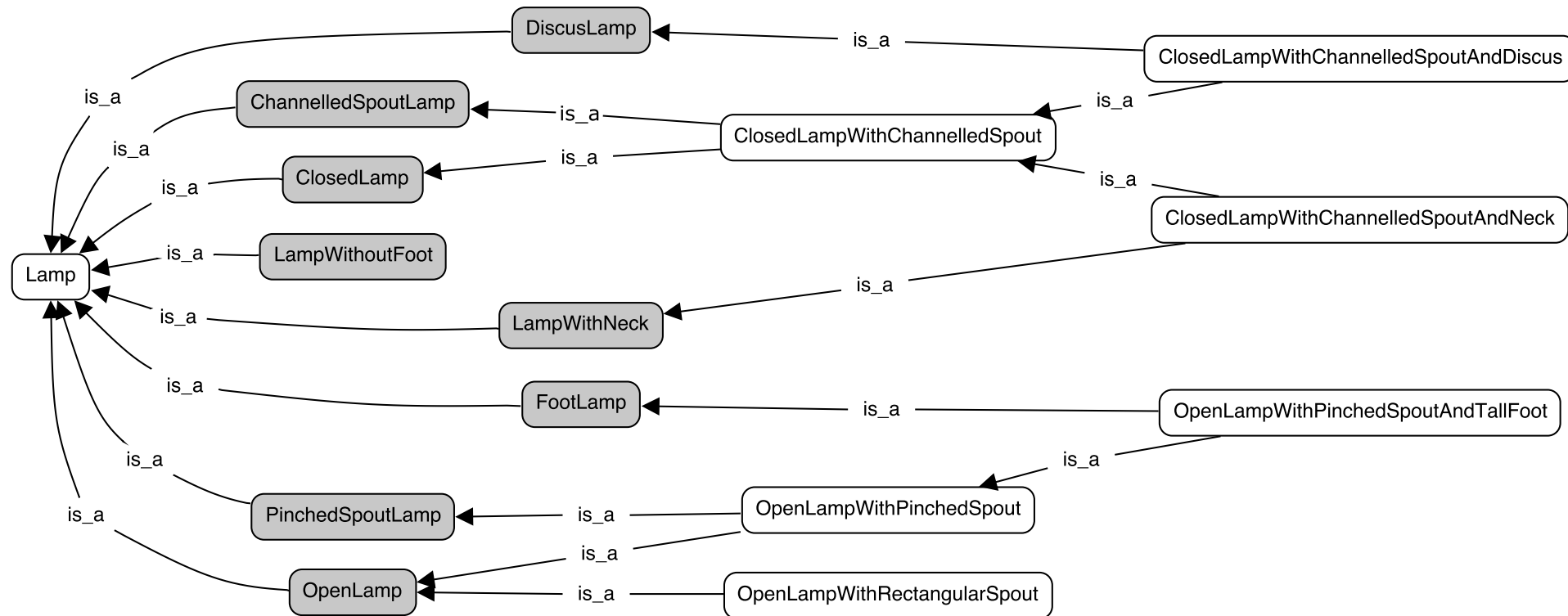
Classification of subtypes of Lamp: asserted hierarchy

- Based on multiple criteria of subdivision: (i) vessel form; (ii) type of spout; (iii) discus or neck; (iv) tall foot



Classification of subtypes of Lamp: inferred hierarchy

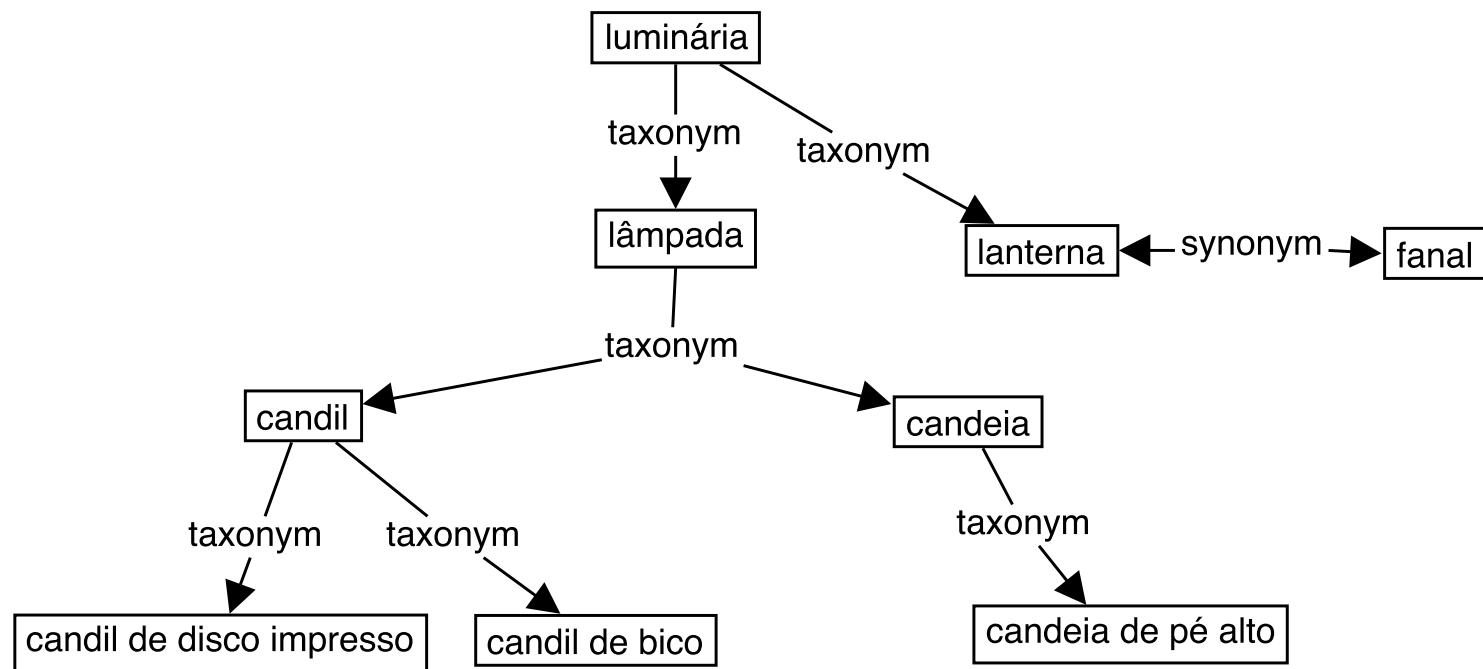
- Inferred with a reasoner within Protégé (Hermit)



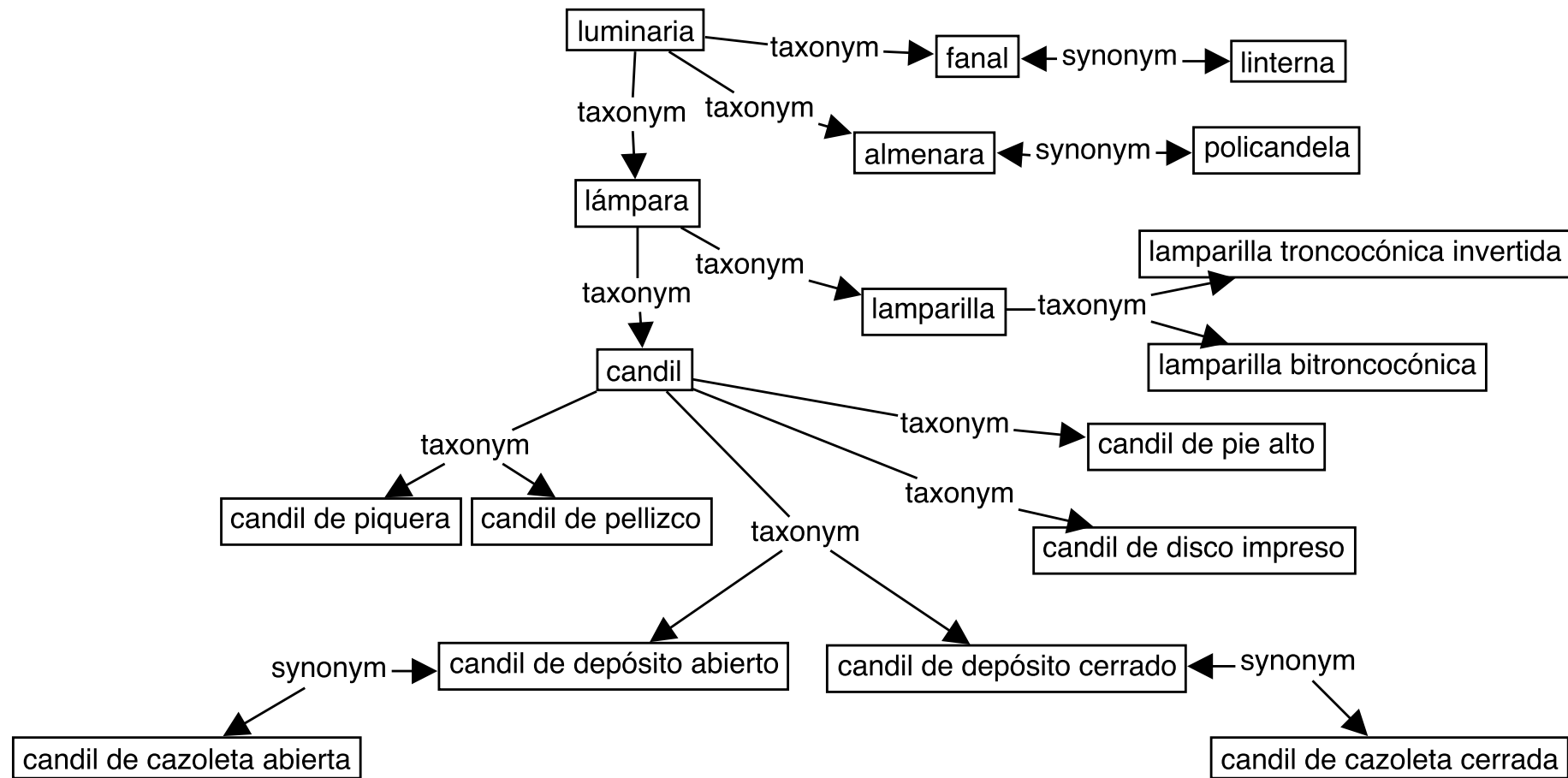
Modelling of artefact terms

- Simple and multiword terms extracted from the bilingual corpus
- Organisation in language-specific lexical networks
 - Taxonomy and synonymy relations between terms
- Lemon as a model for representing grammatical and semantic information about the terms
 - Part of speech, gender, lemma, etc.
 - Relationship between terms and ontological predicates
 - Lexicosemantic relations (language-specific)

Lighting artefact terms in Portuguese



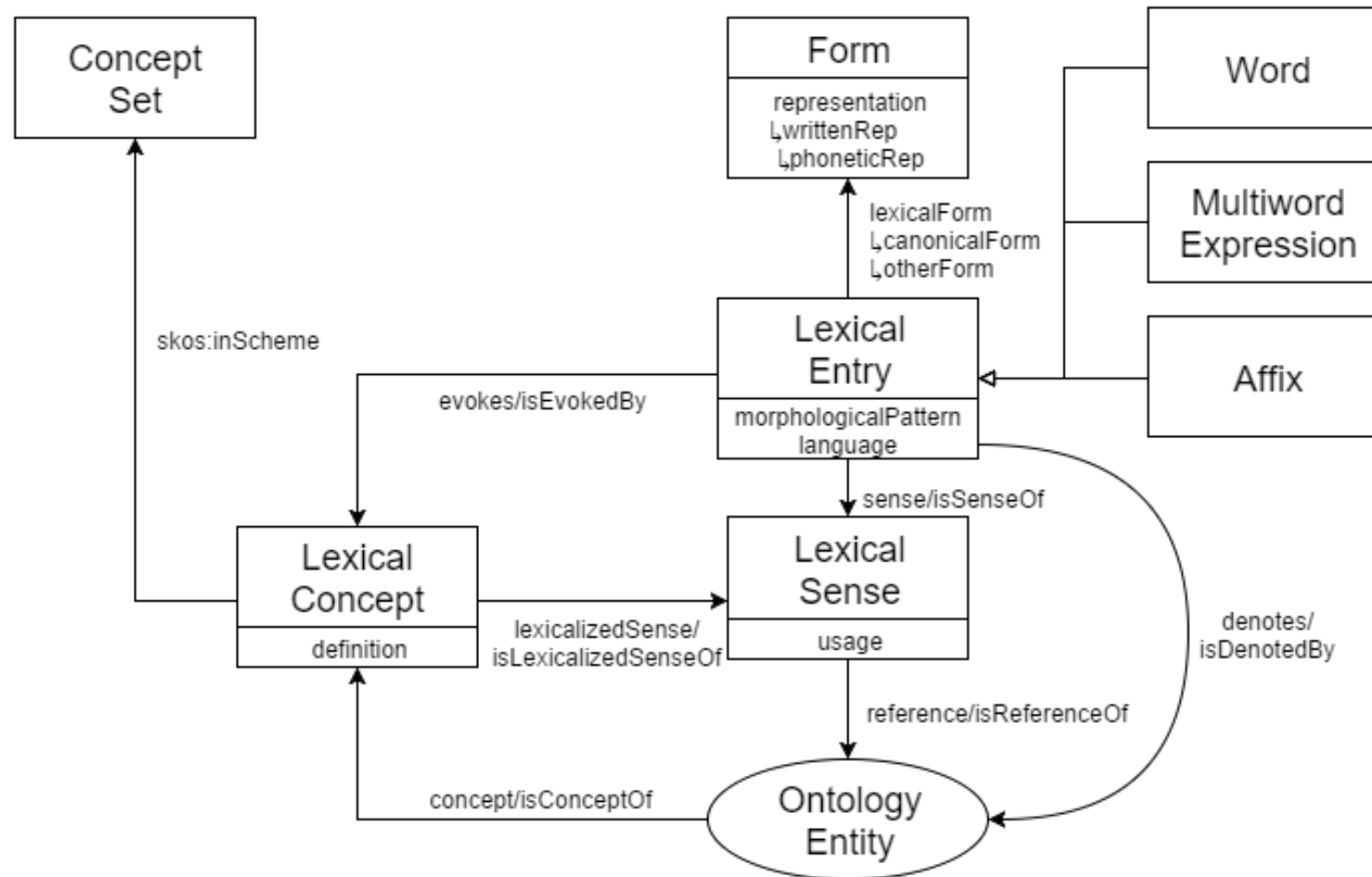
Lighting artefact terms in Spanish



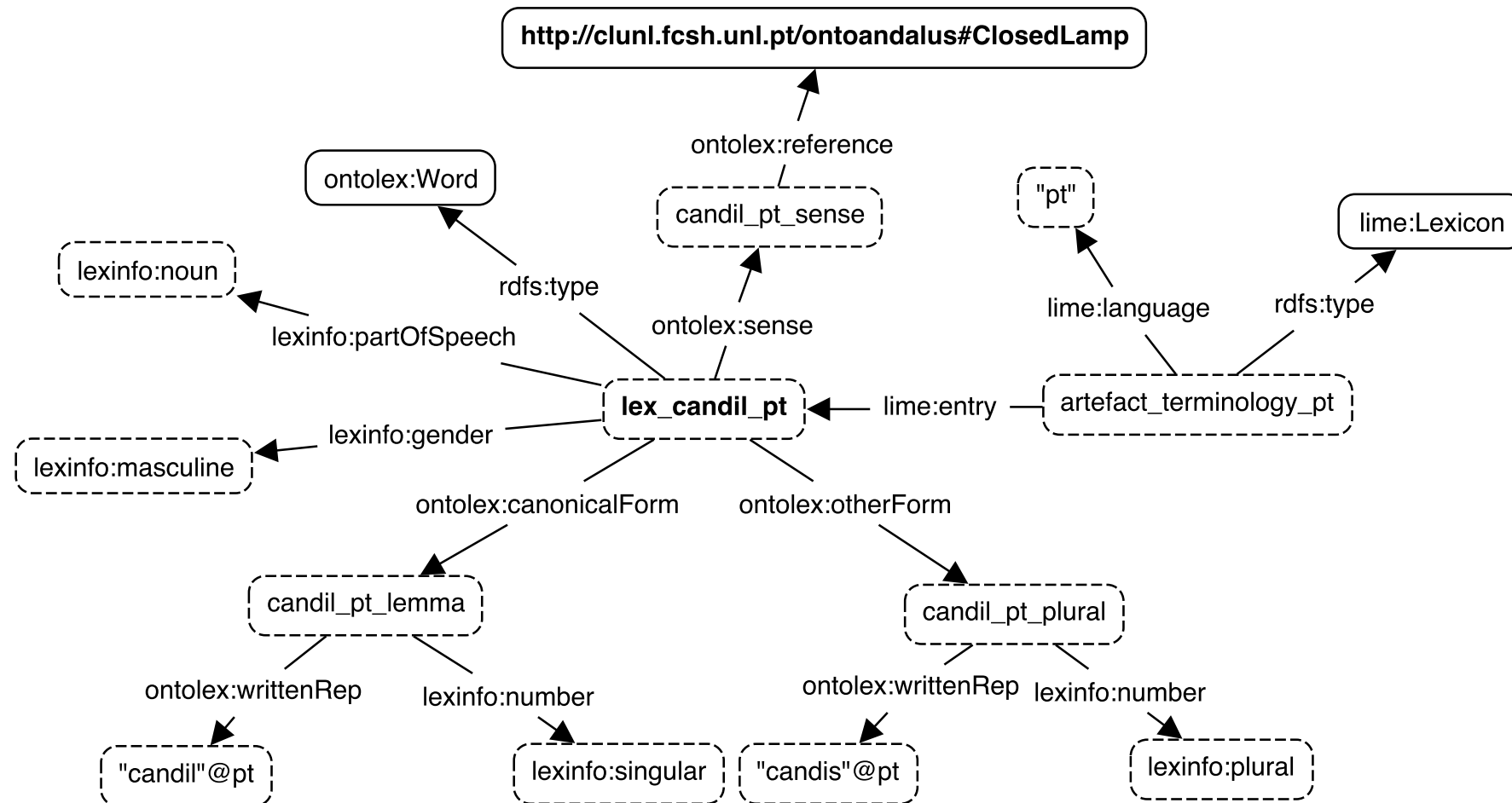
Lemon: Lexicon Model for Ontologies

- Model developed by the W3C Ontology-Lexicon Community Group
 - Enriching ontologies with linguistic information
- Organised in several modules
 - *Ontolex* (core module): for linking an ontology with a lexicon
 - *Synsem*: syntactic and semantic information
 - *Decomp*: decomposition of multiword expressions
 - *Vartrans*: variation and translation
 - *Lime*: linguistic metadata
- *Lexicog*: lexicographic module for describing dictionaries

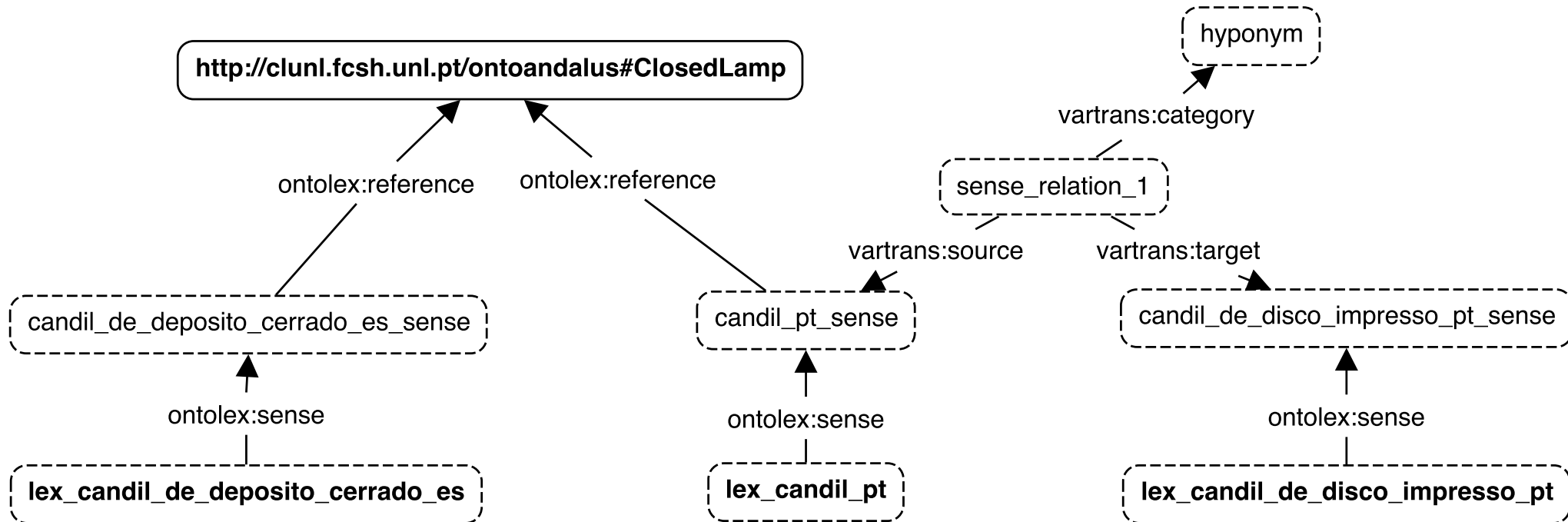
Ontolex: the core module of Lemon



Modelling the Portuguese term *candil*



Lexicosemantic relations: hyponymy and equivalence



In conclusion

- OntoAndalus: elements for knowledge representation in pottery studies of al-Andalus
 - DOLCE+DnS Ultralite allows for a detailed conceptualisation
 - Template for defining/describing artefact types
- Modelling of Portuguese and Spanish terms
 - Lemon allows to draw a relationship between the ontology and lexicon of the domain
 - Representation of rich grammatical and semantic information

Acknowledgement

Research financed by Portuguese National Funding through the FCT – Fundação para a Ciência e Tecnologia as part of the project Centro de Linguística da Universidade NOVA de Lisboa – UID/LIN/03213/2020.

References

- Almeida, B. (2019). *Terminology and knowledge representation: ceramic artefacts of al-Andalus*. Doctoral thesis, Universidade NOVA de Lisboa.
- Almeida, B. and R. Costa (2019). OntoAndalus: an ontology of Islamic artefacts for terminological purposes. Manuscript submitted for publication.
- Almeida, B., R. Costa, and C. Roche (2019). The names of lighting artefacts: extraction and representation of Portuguese and Spanish terms in the archaeology of al-Andalus. *Revue TAL 60(3)*, 113–137.
- Bosque-Gil, J., J. Gracia, J. P. McCrae, P. Cimiano, S. Stolk, F. Khan, K. Depuydt, J. Does, F. Frontini, and I. Kernerman (2019). The OntoLex Lemon Lexicography Module: final community report. Technical report, Ontology-Lexicon Community Group.
- Brin-Henry, F., R. Costa, and S. Desprès (2019). TemPO: towards a conceptualisation of pathology in speech and language therapy. In *Proceedings of the TALN-RECITAL 2019 conference*, Toulouse, pp. 69–80.
- Bugalhão, J., H. Catarino, S. Cavaco, J. Covaneiro, I. C. Fernandes, A. Gomes, S. Gómez Martínez, M. J. Gonçalves, M. Grangé, I. Inácio, G. Lopes, and C. Santos (2010). CIGA: projecto de sistematização para a cerâmica islâmica do Gharb al-ândalus. *Xelb 10*, 455–476.
- Cruse, D. A. (1986). *Lexical semantics*. Cambridge: Cambridge University Press.
- Durán-Muñoz, I. and M. Bautista-Zambrana (2013). Applying ontologies to terminology: advantages and disadvantages. *Hermes 51*, 65–77.
- Gangemi, A. (2016). Dolce+D&S Ultralite and its main ontology design patterns. In P. Hitzler, A. Gangemi, A. Janowicz, A. A. Krisnadhi, and V. Presutti (Eds.), *Ontology engineering with ontology design patterns: foundations and applications*, pp. 81–103. Amsterdam: IOS Press.
- Gómez Martínez, S. (2004). *La cerámica islámica de Mértola: producción y comercio*. Doctoral thesis, Universidad Complutense de Madrid.
- Meyer, I., D. Skuce, L. Bowker, and K. Eck (1992). Towards a new generation of terminological resources: an experiment in building a terminological knowledge base. In *COLING '92 Proceedings of the 14th conference on Computational linguistics*, Volume 3, Stroudsburg, PA, pp. 956–960. Association for Computational Linguistics.
- Musen, M. A. (2015). The Protégé project: a look back and a look forward. *AI Matters 1(4)*, 4–12.

References

- Rice, P. (1987). *Pottery analysis: a sourcebook*. Chicago, IL: The University of Chicago Press.
- Roche, C. (2012). Ontoterminology: how to unify terminology and ontology into a single paradigm. In *LREC 2012*, Madrid, pp. 2626–2630.
- Roche, C., R. Costa, S. Carvalho, and B. Almeida (2019). Knowledge-based terminological e-dictionaries: the EndoTerm and al-Andalus pottery projects. *Terminology* 25(2), 262–294.
- Rosselló-Bordoy, G. (1978). *Ensayo de sistematización de la cerámica árabe en Mallorca*. Palma de Mallorca: Institut d'Estudis Baleàrics.
- Rosselló-Bordoy, G. (1991). *El nombre de las cosas en al-Andalus: una propuesta de terminología cerámica*. Palma de Mallorca: Museo de Mallorca.
- Santos, C. and R. Costa (2015). Domain specificity: semasiological and onomasiological knowledge representation. In H. J. Kockaert and F. Steurs (Eds.), *Handbook of terminology*, Volume 1, pp. 153–179. Amsterdam: John Benjamins.
- Temmerman, R. and K. Kerremans (2003). Termontography: ontology building and the sociocognitive approach to terminology description. In *Proceedings of CIL17*, Prague. Matfyzpress.
- Torres, C., S. Gómez Martínez, and M. B. Ferreira (2003). Os nomes da cerâmica medieval: inventário de termos. In *Actas das 3as Jornadas de Cerâmica Medieval e Pós-Medieval*, Tondela, pp. 125–134. Câmara Municipal de Tondela.
- Vallejo Triano, A. and J. Escudero Aranda (1999). Aportaciones para una tipología de lacerámica común califal de Madinat al-Zahra. *Arqueología y territorio Medieval* 6, 133–176.
- W3C Ontology-Lexica Community Group (2016). Lexicon model for ontologies: community report. Technical report, W3C Ontology-Lexica Community Group.