Terminology and ontology development in the domain of Islamic archaeology[*]

Bruno Almeida,[1] Christophe Roche[2] and Rute Costa[3]

[1] NOVA CLUNL, Faculdade de Ciências Sociais e Humanas, Universidade NOVA de Lisboa, Avenida de Berna, 26-C 1069-061 Lisboa, Portugal
`bruno.r.almeida@gmail.com`
[2] Condillac Research Group, LISTIC, Université de Savoie Mont Blanc, Campus Scientifique, 73376 Le Bourget du Lac, France
`christophe.roche@univ-savoie.fr`
[3] NOVA CLUNL, Faculdade de Ciências Sociais e Humanas, Universidade NOVA de Lisboa, Avenida de Berna, 26-C 1069-061 Lisboa, Portugal
`rute.costa@fcsh.unl.pt`

Abstract. This paper describes an example regarding the terminology of Islamic pottery artefacts in Portuguese and Spanish in the context of an ongoing Ph D project. The approach followed in this paper places knowledge representation at the core of terminology work. More specifically, the development of an ontology, i.e. a formal and computational conceptualisation, enables the integration of a multilingual termbase in the semantic web as linked data, targeted at experts and students of archaeology. This approach allows for the preservation of linguistic diversity, as reflected by the different linguistic practices engaged by Portuguese and Spanish archaeologists in scholarly communication.

Keywords: terminology, knowledge representation, ontologies, multilingual termbases, semantic web, pottery artefacts, Islamic archaeology

1   Terminology and Islamic archaeology: the case of pottery artefacts

Islamic presence in the Iberian Peninsula covered a period of nearly eight centuries (from 711 to 1492 A.D.), and left behind a wide range of materials, such as pottery, architectural fragments, weaponry, jewellery and glassware. For many decades, archaeologists in Portugal and Spain have worked on the description, analysis and comparison of these objects, focusing on properties such as function, shape, materials, manufacturing and decorative techniques.

Pottery is considered to be one of the most important types of artefacts for archaeologists, not only because of its high durability but also due to its cultural significance. According to Kipfer, pottery is "often one of the clearest indicators of cultural differences, relations and developments" (Kipfer, 2000, p. 452). Since the date of manufacture usually can be determined, pottery sherds are also important in dating other finds (*ibid.*). In the last decades, the study of Islamic pottery in Portugal and Spain has furthered the understanding of the culture and society of the al-Andalus:[1] its eating habits, everyday life, trade relations, technical development and even its symbolism and ideology (Gómez Martínez, 2004).

A significant part of this knowledge is only made possible by the typological analysis of pottery artefacts, which enables the comparison and study of related finds. Within archaeology, 'typology' is defined as "the classification of objects, structures, or specimens by subdividing observed populations into a theoretical sequence or series of groups (types) and subgroups (subtypes) according to consideration of their qualitative, quantitative, morphological, formal, technological, and functional attributes." (Darvill, 2009).

In Portugal, the lack of terminology harmonisation has been referred in the past as a hurdle in scholarly communication in the domain of Islamic archaeology (Torres, Gómez Martínez, & Ferreira, 2003). Furthermore, terminology work is seen as a means to acquire and organise expert knowledge in this domain (*ibid.*). In recent years, the need to revitalise the studies on Islamic pottery in Portugal has led to the creation of the CIGA research group (*Cerâmica Islâmica do Gharb al-Ândalus*), which presently consists of twelve archaeologists.[2] The focus of this group was the creation of a shared database describing the most representative instances of Islamic pottery in the Gharb al-Andalus[3] (Bugalhão et al., 2010). Underlying the creation of this database is a common typology and terminology of artefacts, shapes, and manufacturing and decorative techniques. CIGA's typology of artefacts is based on eight classes, according to the theoretical purpose of the objects, namely: (i) storage and transportation, (ii) kitchenware, (iii) tableware, (iv) lighting objects, (v) household objects, (vi) agricultural and handicraft objects, (vii) recreational and ritual objects and (viii) construction materials. Each class is further divided into subclasses according to the formal attributes of the objects. Furthermore, definitions or descriptions in natural language are provided for each subclass, as well as graphical representations in the form of archaeological illustrations (Bugalhão et al., 2010).

The importance of terminology in archaeology, as evidenced by the CIGA group, raises several questions of import to our project, which is centred on the creation of a multilingual termbase in the domain and its integration in the semantic web. In this paper we will focus on the formal and conceptual analysis of Islamic pottery artefacts, following an interdisciplinary approach to terminology. This approach places knowledge representation at the core of terminology work, following previous work

---

[1] 'Al-Andalus' refers to the territory of the Iberian Peninsula and Septimania under Islamic occupation.

[2] More information available at `http://www.camertola.pt/info/ciga`.

[3] Western region of the Iberian Peninsula under Islamic rule, which roughly corresponds to the continental territory of present day Portugal.

in the framework of ontoterminology (Roche, 2007). More specifically, we will show how an ontology may represent a language independent conceptualisation,[4] allowing for the operationalisation of a multilingual termbase meant for experts and students of archaeology.

The example presented in this paper was drawn from the analysis of several texts (quoted below) written by Portuguese and Spanish archaeologists, including relevant graphical information. It should be noted that our conclusions may change as new data is gathered. Translations and equivalent designations in English are provided in order to facilitate communication in this paper.

2   Modelling artefact types: the case of lighting objects in pottery

In the typology of the CIGA group, the class of 'lighting objects' is divided into the subclasses referred to by the Portuguese terms *candil*, *candeia*, *candeia de pé* and *lanterna*. *Candil* is defined as a "lighting object with closed chamber", while *candeia* is defined as a "lighting object with an open chamber".[5] *Candeia de pé* is defined as a "lighting object with an open chamber supported by a high foot". Finally, *lanterna* is described as a "closed form with a globular body and central orifice, used for lighting in open spaces".[6] Fig. 1 illustrates representative instances of the named subclasses of pottery lighting objects.

The available information leads us to infer that *candeia de pé* is actually a subclass of *candeia*, since *candeia de pé* is a lighting object with an open chamber, with the delimiting characteristic of 'being supported by a high foot'.[7] We also infer that the type of object depicted in Fig. 1-II differs from *candeia de pé* by having a flat base instead of a high foot. While *candeia* and *candil* are clearly defined, being distinguished by the configuration of the chamber (open or closed), *lanterna* is described by typical characteristics (i.e. 'globular body with a central orifice') and the more specific purpose of lighting in open spaces. We propose that *candeia* and *candil* should belong to a subclass of lighting objects devised for lighting in closed spaces

---

[4] By 'language independent conceptualisation' we mean a concept system that is not bound by any particular natural language.

[5] Following Rice, an open vessel is generated by an unrestricted orifice, whose "diameter is equal to or greater than the maximum diameter of the body" (Rice, 2015, ch. 13.4.3.1). On the other hand, a closed vessel is generated by a restricted orifice.

[6] According to this information, *candil* can be referred to in English as 'closed lamp', *candeia* as 'open lamp', *candeia de pé* as 'foot lamp' and, finally, *lanterna* as 'lantern'. We should note that these terms may also refer to Islamic artefacts made of other materials besides pottery, which will not be covered in this paper.

[7] According to the ISO terminology standards, a delimiting characteristic is an essential characteristic used for distinguishing between related concepts. By 'essential characteristic' we mean a characteristic that is essential in understanding a concept, which highlights its cognitive nature (ISO 1087-1:2000).

(which we can refer to as 'lamp' in English), since only *lanterna* has the purpose of providing a light source in open spaces.

Fig. 1. Archaeological illustration of the class of 'lighting objects' according to the CIGA group. From left to right: I. *candil*, II, *candeia*, III. *candeia de pé*, IV. *lanterna* (Source: Bugalhão *et al.*, 2010, p. 471).

Regarding the Spanish sources, Rosselló-Bordoy defines *candil* as a "portable or fixed element for domestic lighting" (Rosselló-Bordoy, 1991, p. 174), which corroborates our analysis that the lamp is an object meant for closed spaces. In his earlier work, Rosselló-Bordoy distinguished between several formal variants of *candil*, consisting essentially on the types of artefacts depicted in Fig. 1, I-III (Rosselló-Bordoy, 1978, pp. 48-55). These variants include a subclass referred to in Spanish as *candil de pie alto* (which has an open vessel, similar to Fig. 1, III), four closed variants (depending on the geometrical shape of the chamber), and an open variant without a foot (similar to Fig. 1, II). Therefore, the Spanish term *candil* denotes any type of lamp (open or closed). Rosselló-Bordoy also lists *fanal* or *linterna* within the class of lighting objects, corresponding to the same type of artefact depicted in Fig. 1, IV.[8] Gómez Martínez provides a definition of *fanal* in line with the CIGA group: "*fanal* or *linterna* is defined as a closed form inside which fire is contained for the purpose of lighting in open spaces" (Gómez Martínez, 2004, p. 278). In the case of lamps, Spanish archaeologists also use the terms *candil de piquera* ('nozzled lamp') and *candil de pellizco* ('pinched lamp'). These terms refer to the shape of the beak of these objects, which either have a nozzle or a pinched beak meant for holding a wick. However, 'nozzled lamp' and 'pinched lamp' refer to the same objects as 'closed lamp' and 'open lamp', respectively, as can be observed in the examples represented in Fig. 1. This is evidenced by Navarro Palazón and Jiménez Castillo (2007), who use the terms *candil de piquera* and *candil de pellizco* as synonyms of *candil de cazoleta cerrada* and *candil de cazoleta abierta*, respectively.[9]

---

[8] Rosselló-Bordoy also includes *almenara* in the class of lighting objects, which is defined as "a sort of multiple *candil* or support for holding several *candiles*" (Rosselló-Bordoy, 1991, p. 174). However, this object seems to be ill-defined, as its existence is only documented in metal and not in pottery (Gómez Martínez 2004, p. 277).

[9] These authors also corroborate our analysis that 'foot lamp' is a type of 'open lamp': "Durante los primeros siglos en al-Andalus se empleó un *candil*, denominado genéricamente *de piquera o de cazoleta cerrada*, derivado de las lucernas clásicas. Hacia la segunda mitad del siglo XII llegan a la Península Ibérica, desde el Mediterráneo oriental, dos nuevos tipos, llamados *de cazoleta abierta o pellizco* […] y *de pie alto* […], *este último es básicamente un candil de pellizco dotado de una peana*." [our emphasis] (Navarro Palazón & Jiménez Castillo, 2007, p. 312).

Our analysis leads to the concept system represented in the UML class diagram shown in Fig. 2, based on the principle of genus and specific difference.[10] Concepts are labelled with identifiers in English.[11]
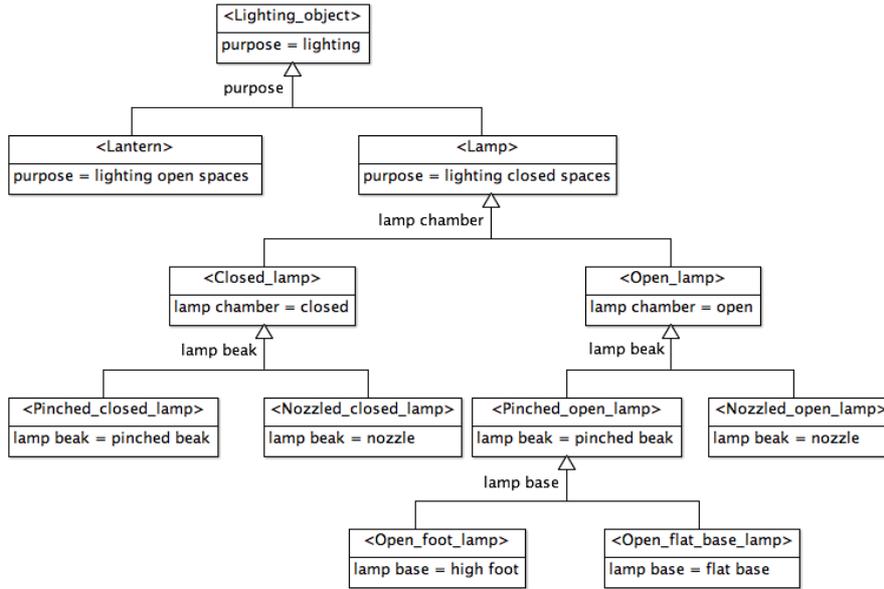

Fig. 2. Concept system of the class of 'lighting objects'.

This concept system can be used as a basis for an ontology of lighting objects. The following axioms provide an ontological definition – i.e. a formal, constructive definition (Roche, 2015) – of the relevant concepts in our ontology:[12]

$$\text{Lighting} \equiv \{\text{lighting\_open\_spaces}\} \sqcup \{\text{lighting\_closed\_spaces}\} \tag{1}$$

$$\text{Lighting\_object} \equiv \text{Islamic\_pottery\_artefact} \sqcap \exists\text{hasPurpose.Lighting} \tag{2}$$

$$\text{Lantern} \equiv \text{Lighting\_object} \sqcap \exists\text{hasPurpose.}\{\text{lighting\_open\_spaces}\} \tag{3}$$

$$\text{Lamp} \equiv \text{Lighting\_object} \sqcap \exists\text{hasPurpose.}\{\text{lighting\_closed\_spaces}\} \tag{4}$$

$$\text{Lighting\_object} \sqsubseteq \text{Lantern} \sqcup \text{Lamp} \tag{5}$$

---

[10] This principle was followed because not only is it consistent with the available data, but also due to its usefulness in producing a conceptualisation in line with the ISO standards on terminology.

[11] Concept identifiers, which are only relevant to identify units of knowledge in a conceptualisation, are represented between angle brackets to further distinguish them from terms (Roche, 2012).

[12] <Pinched_closed_lamp> and <Nozzled_open_lamp> are not defined because they do not have any instances in Islamic pottery. Therefore, in this domain, a <Closed_lamp> is always a <Nozzled_closed_lamp> and an <Open_lamp> is always an <Pinched_open_lamp>.

$$\text{Lantern} \sqcap \text{Lamp} \sqsubseteq \bot \tag{6}$$

$$\text{Lamp\_chamber} \equiv \{\text{open}\} \sqcup \{\text{closed}\} \tag{7}$$

$$\text{Open\_lamp} \equiv \text{Lamp} \sqcap \exists \text{hasLampChamber}.\{\text{open}\} \sqcap \forall \text{hasLampChamber}.\{\text{open}\} \tag{8}$$

$$\text{Closed\_lamp} \equiv \text{Lamp} \sqcap \exists \text{hasLampChamber}.\{\text{closed}\}$$
$$\sqcap \forall \text{hasLampChamber}.\{\text{closed}\} \tag{9}$$

$$\text{Lamp} \sqsubseteq \text{Open\_lamp} \sqcup \text{Closed\_lamp} \tag{10}$$

$$\text{Open\_lamp} \sqcap \text{Closed\_lamp} \sqsubseteq \bot \tag{11}$$

$$\text{Lamp\_beak} \equiv \{\text{pinched}\} \sqcup \{\text{nozzle}\} \tag{12}$$

$$\text{Pinched\_open\_lamp} \equiv \text{Open\_lamp} \sqcap \exists \text{hasLampBeak}.\{\text{pinched}\}$$
$$\sqcap \forall \text{hasLampBeak}.\{\text{pinched}\} \tag{13}$$

$$\text{Nozzled\_closed\_lamp} \equiv \text{Closed\_lamp} \sqcap \exists \text{hasLampBeak}.\{\text{nozzle}\}$$
$$\sqcap \forall \text{hasLampBeak}.\{\text{nozzle}\} \tag{14}$$

$$\text{Lamp\_base} \equiv \{\text{high\_foot}\} \sqcup \{\text{flat\_base}\} \tag{15}$$

$$\text{Open\_foot\_lamp} \equiv \text{Pinched\_open\_lamp} \sqcap \exists \text{hasLampBase}.\{\text{high\_foot}\}$$
$$\sqcap \forall \text{hasLampBase}.\{\text{high\_foot}\} \tag{16}$$

$$\text{Open\_flat\_base\_lamp} \equiv \text{Pinched\_open\_lamp} \sqcap \exists \text{hasLampBase}.\{\text{flat\_base}\}$$
$$\sqcap \forall \text{hasLampBase}.\{\text{flat\_base}\} \tag{17}$$

$$\text{Pinched\_open\_lamp} \sqsubseteq \text{Open\_foot\_lamp} \sqcup \text{Open\_flat\_base\_lamp} \tag{18}$$

$$\text{Open\_foot\_lamp} \sqcap \text{Open\_flat\_base\_lamp} \sqsubseteq \bot \tag{19}$$

Delimiting characteristics are represented by roles whose range is specified by individual values. These values belong to the concepts defined in axioms (1), (7), (12) and (15). The covering and disjointness axioms required for this conceptualisation are defined in (5), (6), (10), (11) and (18), (19).

Ontologies allow for the integration of multilingual resources in the semantic web, functioning as their conceptual and computational underpinning. The question now arises regarding the specificity of each language. This will be addressed in the following chapter.

3 The terminology of lighting objects in Portuguese and Spanish

Although we assume that concepts are extra-linguistic constructs, it does not entail that terminology is independent from the linguistic practices engaged by domain experts in scholarly communication. Terms are determined by cultural and linguistic factors (Lerat, 1995), which makes them more than mere labels for concepts: they are,

in fact, lexical items in their own right, acquiring their status by virtue of their usage and recognition within a specialised community of practice.

Turning our attention to the example at hand, it is clear that there is some difference between both languages. In Portuguese, there does not seem to be a suitable term for <Lamp>, as the archaeologists use the more specific terms *candeia* and *candil*. There is, however, evidence in Portuguese texts that the concepts denoted by these terms are closely related.[13] Nevertheless, <Lamp> should remain an unnamed concept in this language in our termbase in order to reflect the specificity of the linguistic practices of Portuguese archaeologists. Fig. 3 represents the Portuguese terminology of lighting objects according to the data available at this time.[14]
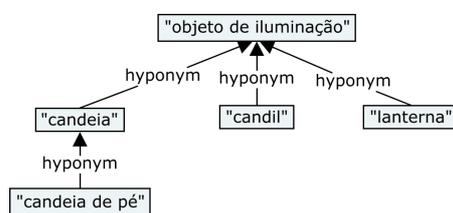
Fig. 3. The terminology of lighting objects in Portuguese.

In the case of Spanish, every concept is denoted by at least one term in scholarly communication, including three notable cases of synonymy. The information regarding the Spanish terminology is represented in Fig. 4.
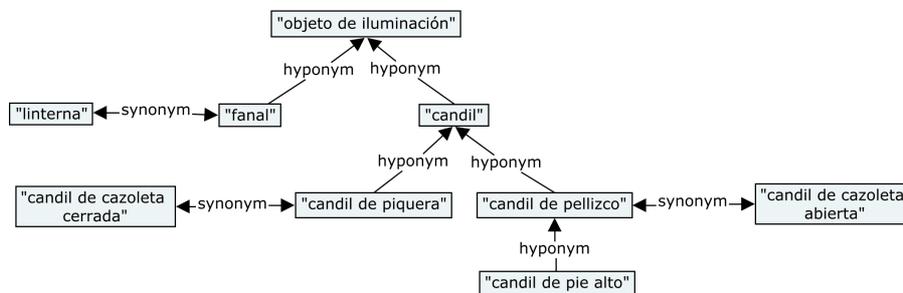
Fig. 4. The terminology of lighting objects in Spanish.

As we can see, neither of these lexical networks, which represent language specific information, is isomorphic to the concept system outlined in Fig. 2, which represents knowledge shared within a community of practice. From an onomasiological point of

---

[13] For instance: "[…] distinguindo deste modo o CANDIL, de forma fechada, da CANDEIA que seria a forma aberta que se manteve praticamente até aos nossos dias" [emphasis in the original] (Torres, Gómez Martínez & Ferreira, 2003, p. 129).

[14] Terms are represented between double quotation marks (Roche 2012). The lexical networks in this section are based on the relations of hyponymy (generic-specific relation between term meanings) and synonymy (relation of equivalence between term meanings).

view, the concepts in our ontology are denoted by the following terms in each language:

- <Lighting_object> *isDenotedBy* "objeto de iluminação" (pt), "objeto de iluminación" (es);
- <Lantern> *isDenotedBy* "lanterna" (pt), "fanal" (es), "linterna" (es);
- <Lamp> *isDenotedBy* "candil" (es);
- <Closed_lamp> (and <Nozzled_closed_lamp>) *isDenotedBy* "candil" (pt), "candil de piquera" (es), "candil de cazoleta cerrada" (es);
- <Open_lamp> (and <Pinched_open_lamp>, <Open_flat_base_lamp>) *isDenotedBy* "candeia" (pt), "candil de pellizco" (es), "candil de cazoleta abierta" (es);
- <Open_foot_lamp> *isDenotedBy* "candeia de pé" (pt), "candil de pie alto" (es).

The interface between our termbase and the ontology described in the last section can be achieved by adapting a model such as Lemon (Lexicon Model for Ontologies), which is under development by the W3C Ontology-Lexica Community Group.[15] This also facilitates the access to the termbase as linked data in RDF (Resource Description Framework). To give an example, the following RDF code in Turtle syntax represents a terminological entry for "candil" in Portuguese, referring to the concept <Nozzled_closed_lamp>:[16]

```
:candil-pt a ontolex:LexicalEntry, ontolex:Word ;
  ontolex:canonicalForm :candil-pt#CanonicalForm ;
  rdfs:label "candil"@pt ;
  ontolex:language "pt" ;
  ontolex:sense  :candil-pt#Sense .

:candil-pt#CanonicalForm a ontolex:Form ;
  ontolex:writtenRep "candil"@pt .

:candil-pt#Sense a ontolex:LexicalSense ;
  ontolex:reference <http://…/Nozzled_closed_lamp> ;
  skos:definition "Objeto cerâmico de origem islâmica para iluminação doméstica com depósito fechado e bico de canal."@pt .

:senseRelation a vartrans:SenseRelation ;
  vartrans:source :objeto_de_iluminacao-pt#Sense ;
  vartrans:target :candil-pt#Sense ;
  vartrans:category :hyponym .
```

This approach enables the full integration of an archaeology termbase in the semantic web. This facilitates the access to, and manipulation of, terminological and

---

[15] More information available at `https://www.w3.org/community/ontolex/`.

[16] We should note that in this example "candil" refers only to a subclass of pottery artefacts. However, further senses of the term can be defined in order to provide a more complete account of its meaning within Islamic archaeology.

conceptual data by both human and machine agents, which is paramount in the context of information society, allowing for a more efficient construal of knowledge.

4   Concluding remarks

The analysis outlined in this paper is only possible by following an interdisciplinary approach to terminology, looking beyond linguistics and specialised lexicography. This has an important precedent in the work of Wüster (1979), for whom terminology theory overlaps with logic, ontology and information science.

Terminology as a domain emerges from the interaction between disciplines centred on the study of language and knowledge, from which it derives its principles and methods as a discipline. Presently, the object of study of terminology is recognised as being multidimensional and, therefore, irreducible to any particular discipline (Cabré, 2000). We assume that terminology has fundamentally a double dimension: linguistic and conceptual (Costa, 2013; Roche 2015; Santos & Costa, 2015). While the linguistic dimension pertains to terms, their behaviour in discourse and their role within specialised communities of practice, the conceptual dimension consists on the knowledge shared within these communities and how it can be represented for multiple applications (computational or otherwise). Indeed, the core elements of terminology remain the concept (unit of knowledge), the term (specialised lexical item), and the relationship between these elements, in which lies the specificity of terminology as a domain at the crossroads between language and knowledge (Costa, 2013).

In the past decades, terminology has been characterised by a "plurality of theoretical approaches" (Costa, 2006) in which linguistics plays an increasingly dominant role and thereby relegating terminology to a sort of specialised lexicography. However, the need for the operationalisation of multilingual terminology resources, i.e. their computational representation, requires an approach in line with knowledge representation, a field of artificial intelligence, which once again brings into question the need to widen the scope of terminology as an interdisciplinary domain. This opens up important applications for the discipline in the context of information society, from computer assisted translation to SEO, semantic search engines and interactive navigation tools in data repositories (Roche, 2015).

This paper focused on the conceptual dimension of terminology. We saw how lexical networks, which represent language specific information, are not isomorphic to a concept system, which represents shared knowledge in the domain. Placing ontology development at the core of terminology work enables the operationalisation of multilingual terminologies in the semantic web, allowing for the description of the linguistic diversity manifested in scholarly communication.

References

Bugalhão, J. *et al.* (2010). CIGA: projecto de sistematização para a cerâmica islâmica do Gharb al-Ândalus. *Xelb, 10,* 455-476.

Cabré, M. T. (2000). Terminologie et linguistique : la théorie des portes. *Terminologies nouvelles*, *21*, 10-15.

Costa, R. (2006). Plurality of theoretical approaches to terminology. In: H. Picht (Ed.), *Modern approaches to terminological theories and applications* (pp. 79-89). Bern: Peter Lang.

Costa, R. (2013). Terminology and specialised lexicography: two complementary domains. *Lexicographica*, *29*, 29-42.

Darvill, T. (2009). *The concise Oxford dictionary of archaeology* (Online ed.). Oxford: Oxford University Press.

Gómez Martínez, S. (2004). *La cerámica islámica de Mértola: producción y comercio.* Madrid: Universidad Complutense.

ISO 1087-1 (2000). Terminology work – Vocabulary – Part 1: Theory and application. Geneva: ISO.

Kipfer, B. A. (2000). *Encyclopedic dictionary of archaeology.* New York: Springer Science+Business Media.

Lerat, P. (1995). *Les langues spécialisées.* Paris: Presses universitaires de France.

Navarro Palazón, J. & Jiménez Castillo, P. (2007). *Siyasa: estudio arqueológico del despoblado andalusí (ss. XI-XIII).* Granada: Escuela de Estudios Árabes de Granada.

Rice, P. M. (2015). *Pottery analysis: a sourcebook.* (Second ed.). (Kindle ed.). Chicago: The University of Chicago Press.

Roche, C. (2012). Ontoterminology: how to unify terminology and ontology into a single paradigm. In: *Proceedings of the 8th International Conference on Language Resources and Evaluation (LREC-2012)* (pp. 2626–2630). Paris: ELRA.

Roche, C. (2007). Terme et concept : fondements pour une ontoterminologie. In: TOTh 2007 (pp. 1-22). Annecy: Institut Porphyre.

Roche, C. (2015). Ontological definition. In: H. J. Kockaert & F. Steurs (Eds.), *Handbook of Terminology: vol. 1* (pp. 128–152). Amsterdam: John Benjamins Publishing Company.

Santos, C. & Costa, R. (2015). Domain specificity: semasiological and onomasiological knowledge representation. In: H. J. Kockaert & F. Steurs (Eds.), *Handbook of Terminology: vol. 1* (pp. 153–179). Amsterdam: John Benjamins Publishing Company.

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-Ándalus: una propuesta de terminología cerámica.* Palma de Mallorca: Museo de Mallorca.

Torres, C., Gómez Martínez, S., & Ferreira, M. B. (2003). Os nomes da cerâmica medieval: inventário de termos. In: *Actas das 3as Jornadas de Cerâmica Medieval e Pós-Medieval* (pp. 125-134). Tondela: Câmara Municipal.

Wüster, E. (1979). *Introduction to the General Theory of Terminology and Terminological Lexicography*. Vienna: Springer.