Linguistic realization of conceptual features in terminographic dictionary definitions

Esperanza Valero
TecnoLeTTra Team
Universitat Jaume I
Castellón, Spain
mvalero@trad.uji.es

Amparo Alcina
TecnoLeTTra Team
Universitat Jaume I
Castellón, Spain
alcina@trad.uji.es

Abstract
In the knowledge society, researchers on lexicography recognize the need to advance in the structure of dictionary information so that it can be understood by people and computers. In relation with dictionary definitions, the adoption of models or templates would be advantageous for the generation of complete definitions and also for the extraction of semantic information from them.

In this study, we manually analyzed specialized dictionary definitions of the ceramic field belonging to the conceptual groups of ceramic processes and ceramic defects in order to identify the relevant conceptual features, such as physical aspect or function, and the linguistic realization of these features in the definitions.

Results can help to extract information from definitions and to also generate formalized dictionary definitions.

Keywords
Terminographic definition, features, linguistic markers, patterns, conceptual information

1. Introduction
Definitions are a very valuable source of semantic information for different tasks such as the creation of ontologies, lexicography, terminology and natural language processing. Applied research into computational lexicography and terminology is being carried out to recognize definitions in specialized texts and to analyze the semantic relationships that occur in these definitions [19, 20, 14, 17]. Dictionary definitions are other means of obtaining semantic data and of discovering relationships between the concepts of a domain. We can find numerous studies intended to extract information from the definitions in existing dictionaries. Most of them aim to extract taxonomic relationships [5, 11]. However, more effort is needed to exploit the rest of the information present in the definitions, such as specific features and non-hierarchical conceptual relationships [4, 6, 7, 10]. When attempting to extract this information, the biggest problem that automatic systems face is the lack of homogeneity and systematicity in definitions. Most dictionary definitions present inconsistent and incomplete information as well as terms which should be equally treated and which are defined in a very different way [9]. Many authors highlight the need to create more standardized, precise definitions in a format that is understandable for computers and humans alike so that extraction and reusability are easier.

One of the objectives of the ONTODIC project is to develop a system to assist the terminographer in the elaboration of definitions. This system will semi-automatically generate definitions based on a definitional template and a domain ontology. The definitional template will contain the necessary linguistic markers to introduce each feature into the definition. An example of this definitional template for the conceptual group ceramic tiles is proposed by Alcina [1]:

“A ceramic tile whose shape is X and size is Y, and is decorated with Z to serve as Q”

Variables x, y, z and q will be replaced by the values that each feature in the concept description acquires.

In this study, our objective is to observe the type of features which are relevant for the description of two conceptual groups (ceramic processes and ceramic defects) in this domain and to analyze how they are expressed linguistically in the 222 definitions taken from three specialized dictionaries of the ceramic field. We studied the linguistic patterns used in the definitions to introduce features. This set of features and their linguistic realization can, on the one hand, be useful to extract information from dictionaries and, on the other, to generate definitions.

2. Semantic information extraction from texts and dictionary definitions
Several studies have been conducted to automatically identify the related concepts in a corpus. However, most of them focus on the English and French languages with few centering on Spanish [23, 19]. The authors agree that a good way to extract information from texts is by searching for recurrent patterns.

According to [15], knowledge patterns are “words, word combinations or paralinguistic features which frequently indicate conceptual relations”. The authors distinguish three types of patterns:

---

1 ONTODIC: Methodology and Technology for the creation of onomasiologic dictionaries based on ontologies. Terminological resources for the e-translation, financed by the Education and Science Ministry (TSI2006-01911).
Lexical patterns which are words that indicate a relation. For example, is a can indicate hypernymy.

Grammatical patterns which are combinations of part-of-speech. The NOUN+VERB pattern can indicate the function relation as in the sentence: L’unité centrale effectue le traitement.

Paralinguistic patterns which include punctuation, parenthesis, text structure, etc. The authors provide the following example where the questions introduce the hypernymic relation. Qu’est-ce qu’un réseau ? Un réseau est un ensemble de ressources à la […]

According to [11], “definitions use a sublanguage of natural language”. Dictionary definitions have a more explicit semantic language structure which is accessible for analysis [24].

Conceptually, analytical definitions have two main components: genus which reflects the hierarchical conceptual organization in the specialized domain in the definitions and differentiae or specific features which distinguish the defined concept from their co-hyponyms and reflect all kinds of relations in a domain.

In [24] it is acknowledged that it is less difficult to extract generic terms from definitions for computational semantics than to extract specific features. It is not possible to find a complete list of the semantic information that can be extracted from definitions in the literature. Sager and L’Homme [18] considered that this part of the definitions is not susceptible to codification following a restricted set of features and that this would be an important step for the systematization of terminographic definitions.

In order to find regularities among the definitions, we restricted the analysis to the conceptual groups of ceramic processes and ceramic defects which are described by a limited set of features.

3. Methodology

In this study, we use dictionary definitions to extract the features and relations of concepts. The aim of this study is to discover linguistic patterns that are used in the definitions to denote the specific features of concepts. In order to find regularities among the definitions, we restricted the analysis to the conceptual groups of ceramic processes and ceramic defects which are described by a limited set of features.

3.1 Definition selection

We analyzed 222 definitions from three dictionaries of the ceramics domain: Diccionario científico-práctico de la cerámica, Diccionario de cerámica y Terminología de los defectos cerámicos [12, 13, 21]. These dictionaries are published on paper and have been digitalized [3].

From these dictionaries, we extracted the definitions of the concepts included in the categories of ceramic production processes and defects in the ceramic product. It is important for our study that the conceptual groups to be homogeneous enough to be able to identify a relevant set of features for all of them. In total, we analyzed 135 definitions of the conceptual group ceramic processes and 87 ceramic defects.

3.2 Identification of conceptual features in the definitions

The definitions in these dictionaries vary considerably as regards to the use of words and their formal structure [2]. Many of them follow the analytical model with the formula Definiendum = genus + differentia, although we can also find descriptions by means of synonyms, paraphrases, etc. In this analysis, we did not focus on the formal aspects of the definitions, but on the conceptual features that they provide. We analyzed the definitions in order to obtain a set of features which are commonly used in the descriptions of the concepts of each conceptual group. We followed the proposal of [16] which distinguishes between the name of the feature and its value. The name of the feature acts as a label that indicates the content of the feature, while the value offers specific information about a concept. For example, cause is the name and friction the value of a feature for the concept abrasion.

This conceptual analysis was carried out by segmenting the information obtained from the definition and by assigning a label or code which describes the type of information that each fragment represents, as seen in Figure 1. In order to carry out this analysis, we used the program for the qualitative analysis named Atlas.ti. This program allows us to: segment the information, assign a descriptive code, create relations between them, obtain graphic representations of the conceptual structure of data and query as in a database.

The result is a list of essential features to describe each category and a set of values for each of these features. The features detected for these categories are as follows (frequency of appearance in the corpus is indicated in the parenthesis):

- **Production process features**: PROCEDURE (69), OBJECTIVE (103), PATIENT (56), MATERIAL STATE (15), INSTRUMENTS (22), PREVIOUS STAGE (6) and NEXT STAGE (8).

- **Ceramic defect features**: PHYSICAL ASPECT (54), ZONE (16), CAUSE (44), PHASE (15), METHOD (5) and PRODUCT (25).
Figure 1. Identifying conceptual features

lágrima: Defecto de aplicación de los esmaltes [STAGE], consistente en gotones [PHYSICAL ASPECT] que aparecen en la superficie de la capa [ZONE] por haberla depositado incorrectamente, ya sea por usar pincel de punta en vez de pinceleta de punta chata, o por falla en la boquilla del aerógrafo si se esmaltó por pulverización mediante compresor [CAUSE].

Teardrop: defect in the glazing process [STAGE] consisting of drops [PHYSICAL ASPECT] that appear on the surface layer [ZONE] due to incorrect application, either through using a pointed brush instead of a flat brush or because of a fault in the airbrush nozzle if glazing was applied by spray [CAUSE].

hinchamiento: Aumento en las dimensiones o volumen aparente [PHYSICAL ASPECT] de un artículo [PRODUCT] causado por la reacción con el agua o el vapor de agua [CAUSE].

Bloating: Increase in the dimensions or bulk of the volume [PHYSICAL ASPECT] of an object [PRODUCT] caused by a reaction to water or water vapor [CAUSE].

3.3 Identification of linguistic markers

We marked the linguistic expression that precedes each feature. In some cases it is possible to identify a linguistic marker that introduces a particular feature. As Figure 2 illustrates, the feature OBJECTIVE of a process is introduced into the first definition by the linguistic pattern “que sirve para” (in English “which serves to”) and in the second, by the marker “con el fin de” (“for the purpose of”).

Figure 2. Examples of linguistic patterns in the definitions

cocción de decoración: cocción que sirve para madurar los efectos decorativos aplicados previamente a las pieza [OBJECTIVE].

decoration firing: firing which serves to mature the decorative effects previously applied to the piece [OBJECTIVE].

compactación de polvos: Operación de condensar los materiales pulverulentos con el fin de obtener productos con la mayor densidad posible [OBJECTIVE].

powder compaction: Operation to condense the powdery materials with the purpose of obtaining products that are as dense as possible [OBJECTIVE].

cocción lenta: cocción que se desarrolla a una velocidad más pequeña de la acostumbrada [PROCEDURE].

slow firing: firing that is done at a lower speed than usual [PROCEDURE].

In many cases, however, it was not possible to identify a linguistic marker to introduce the feature, rather recurrent syntactic structures in the expression of the feature itself. As we can see in Figure 3, the feature PROCEDURE is expressed in both definitions with a sentence in which the main verb is in the gerund.

Figure 3. Recurrent syntactic structures in definitions
colar: Formar una pieza vertiendo un líquido o una masa plástica en un molde [PROCEDURE], por fraguado o enfriamiento.
to cast: To form a tile by pouring a liquid or a plastic mass into a mold [PROCEDURE] by setting or cooling.

aclarar, diluir (un color): Rebajar la intensidad de un color añadiéndole componente blanco o agente blanqueante [PROCEDURE].
to lighten, to dilute (a color): To reduce the intensity of a color by adding a white component or a bleaching agent [PROCEDURE].

4. Results

The results obtained in this analysis consist in a set of linguistic markers and recurrent syntactic structures denoting different types of features in the definitions. Linguistic markers introduce the feature; one example would be en forma de (Eng. as ……-like) to describe the physical aspect of a defect. Recurrent syntactic patterns are syntactic constructions that mark a particular meaning; for example, the use of gerund constructions to express the mode of carrying out a process.

Table 1: Linguistic markers in the features of ceramic defects

| Pattern | Examples from corpora |
|---------|-----------------------|
| **PHYSICAL ASPECT** | |
| aparece (2) | aparece rodeada de una envoltura vítrea, acompañada generalmente de una cuerda en forma de cola |
| en forma de (2) | en forma de motas o agujeritos |
| se parece* a/que parece (2) | se parecen a "picaditas de alfiler" |
| caracterizado por (2) | caracterizado por una aspereza extrema |
| presenta (1) | presenta pequeñas arrugas u olas |
| que se manifiesta (1) | que se manifiesta en forma de huecos o pequeñas burbujas reventadas |

| Pattern | Examples from corpora |
|---------|-----------------------|
| **PROCEDURE** | |
| en forma de (2) | en forma de motas o agujeritos |
| se parece* a/que parece (2) | se parecen a "picaditas de alfiler" |
| caracterizado por (2) | caracterizado por una aspereza extrema |
| presenta (1) | presenta pequeñas arrugas u olas |
| que se manifiesta (1) | que se manifiesta en forma de huecos o pequeñas burbujas reventadas |
### CAUSE

| Cause                                                                 | Translation                                           |
|---------------------------------------------------------------------|-------------------------------------------------------|
| `cuando+oración 8)`                                                  | when clause                                          |
| `se deb* a (7)`                                                     | it is due to                                         |
| `debido a (5)`                                                      | due to                                               |
| `causad* por / porque (4)`                                         | caused by / caused because of                        |
| `por+ SN (3)`                                                       | because of +NP                                       |
| `produida por (2)`                                                  | produced by                                          |
| `por causa de (1)`                                                 | because of                                           |
| `como consecuencia de (1)`                                         | as a result of                                        |
| `formad * por (1)`                                                  | formed by                                            |
| `originadas por (1)`                                               | caused by                                            |
| `procedentes de (1)`                                               | from                                                 |
| `surgida por (1)`                                                  | having emerged by                                   |
| `da lugar a (1)`                                                   | gives rise to                                        |
| `por efecto de (1)`                                                | as a result of                                        |
| `resultado de (1)`                                                 | as a result of                                        |
| `consiste en (1)`                                                  | it consists of                                       |
| `cuando no está bien compensada su fórmula`                         | when not properly compensated its formula           |
| `se debe a que el coeficiente de dilatación de la pasta es más elevado que el de esmalte` | it is due to                                          |
| `debido a defectos del secado o de la pasta`                        | due to defects of the drying or of the glaze         |
| `causada por fallo mecánico debido a la tensión.`                   | caused by mechanical failure due to tension          |
| `por enfriamiento demasiado rápido`                                 | caused by enervating excessively fast                |
| `por producirse ésta de manera irregular`                           | caused by producing this in an irregular manner      |
| `produída por un desprendimiento gaseoso anormal`                   | produced by a gaseous deslagging abnormally          |
| `por causa del enfriamiento demasiado rápido post cocción,`         | caused by excessive cooling after drying,          |
| `como consecuencia de su disolución parcial`                        | as a result of its dissolution partial               |
| `formada generalmente por la penetración de vidrio entre las partes del molde` | formed generally by the penetration of glass between the parts of the mold |
| `originadas por impurezas depositadas sobre el vidrio caliente durante su trabajo` | caused by impurities deposited on hot glass during its treatment |
| `procedente de herrumbre o de otras impurezas`                      | derived from rust or other impurities                |
| `surgida por retracción del vidrio`                                 | caused by contraction of glass                       |
| `La presencia de carbón [...] da lugar al “corazón negro”`          | carbon presence [...] causes the “black heart”       |
| `por efecto de un recalentamiento excesivo o de la acción de gases`  | as a result of excessive reheating or gas action     |
| `resultado de haber tomado un objeto una forma convexa`             | result of having taken an object a convex shape      |
| `Consiste en que una pieza ha perdido fragmentos del vidriado.`     | It consists of that a piece has lost fragments of glass. |

### STAGE

| Stage                                                                 | Translation                                           |
|---------------------------------------------------------------------|-------------------------------------------------------|
| `durante+SN (7)`                                                     | during + NP                                           |
| `después de + SN(2)`                                                | after +NP                                             |
| `originada (1)`                                                     | caused during                                         |
| `en (1)`                                                            | which starts in                                        |
| `en el secado`                                                       | in the drying                                         |
| `al levantar el vidrio`                                             | by lifting the glass                                  |
| `Por vía seca`                                                       | By means of dry                                      |
| `que pueden sufrir los esmaltes`                                    | that can suffer the enamels                            |
| `en particular/es (2)`                                             | in particular/an                                        |
| `En determinadas producciones cerámicas`                            | In certain ceramic productions                         |
| `surgida en (1)`                                                    | emerged in                                            |
| `surgida en el vidrio`                                              | caused in the glass                                   |
| `en la superficie`                                                  | in the surface                                        |
| `defecto de la superficie`                                           | defect of the surface                                 |
| `Aparece* (2)`                                                      | Appears                                               |
| `aparece en la superficie de los productos cerámicos`               | appears in the surface of the ceramic products        |
| `Situada (1)`                                                       | Located                                               |
| `Situada en la cara interna`                                        | Located in the inner face                             |
| `Próxima a (1)`                                                     | Near                                                   |
| `Próxima a la superficie`                                           | Close to the surface                                  |

### METHOD

| Method                                                                 | Translation                                           |
|---------------------------------------------------------------------|-------------------------------------------------------|
| `cuando+oración (2)`                                                | when clause                                          |
| `cuando se esmalta por baño el interior de los jarrones`             | when glazed in the interior of the jars              |
| `Por vía seca`                                                       | By means of dry                                      |
| `que suelen presentar los esmaltes cerámicos`                        | that usually present ceramic enamels                  |
| `En determinadas producciones cerámicas`                            | In certain ceramic productions                         |
| `surgida en (1)`                                                    | emerged in                                            |

### ZONE

| Zone                                                                 | Translation                                           |
|---------------------------------------------------------------------|-------------------------------------------------------|
| `en +SN (8)`                                                        | in +NP                                                |
| `de+SN (3)`                                                         | of +NP                                                |
| `Aparece* (2)`                                                      | Appears                                               |
| `aparece en la superficie de los productos cerámicos`               | appears in the surface of the ceramic products        |
| `Situada (1)`                                                       | Located                                               |
| `Situada en la cara interna`                                        | Located in the inner face                             |
| `Próxima a (1)`                                                     | Near                                                   |
| `Próxima a la superficie`                                           | Close to the surface                                  |
### Table 2. Recurrent syntactical structures in the features of ceramic defects

| Feature          | Pattern | Examples from corpora |
|------------------|---------|-----------------------|
| **PHYSICAL ASPECT** | NP (40) | Velo de burbujas finas |
| **ZONE**         | Adjective (1) | superficial |

### Table 3. Linguistic markers in the features of ceramic processes

| Pattern                           | Examples from corpora                                                                 |
|-----------------------------------|--------------------------------------------------------------------------------------|
| **OBJECTIVE**                     |                                                                                        |
| **para** +inf (15) in order to +inf | para dotarle de las propiedades adecuadas.                                          |
| **consistent** +en (2) consisting of | consistente en eliminar de las piezas moldeadas las rebabas y otras protuberancias |
| **que sirve** para (1) / it is for | que sirve para madurar los efectos decorativos                                       |
| **PROCEDURE**                     |                                                                                        |
| **dirigido** +a (1) whose aim is to | Acción dirigida a que un esmalte que da superficies brillantes se transforme…         |
| **por** +SN (14) by+NP            | por inmersión de la pieza en un baño de esmalte                                      |
| **Mediante** (3) by means of       | Mediante su inmersión en un baño de esmalte                                          |
| **que se logra** (1) which is achieved by | Que se logra pasando la mezcla a través del medio adecuado.                          |
| **se desarrolla** +SN (1) is done | que se desarrolla a una velocidad más pequeña de la acostumbrada.                   |
| **se lleva a cabo** +SN (1) is carried out | se lleva a cabo el paso del material a través del tamiz.                            |
| **PREVIOUS PHASE**                |                                                                                        |
| **después de** (3) after           | después de haber sido esmaltada                                                       |
| **NEXT PHASE**                    |                                                                                        |
| **previo/o** (1) previous          | previa homogeneización                                                               |
| **antes de** (3) before            | antes de ser esmaltada                                                               |
| **INSTRUMENTS**                   |                                                                                        |
| mediante (5)                      | mediante un aerógrafo.                                                                |
| **con** +SN (2)                   | Con piedra o herramienta de acero                                                    |
| With +NP                          |                                                                                        |
| **en** +SN (2)                    | en un molde                                                                           |
| **in** +NP                        |                                                                                        |
| **por** +SN (2)/ by+NP            | Por prensas mecánicas                                                                |
| **empleando** (1) using            | empleando molinos de rodillos, bolas o guijarros                                     |
| **haciendo uso de** using (1)     | haciendo uso de un dispositivo (pistola)                                            |
| **se utilizan** +SN (1) are used   | se utilizan resistencias eléctricas                                                  |
| **con ayuda de** (1) the help of   | con ayuda de un calibre o plantilla                                                  |
| **mediante el uso** de (1) by using| mediante el uso de resinas sintéticas de intercambio iónico.                        |
| **por medio de** (1) by means of   | por medio de un pistón                                                               |
| **se requiere**/en (1) is required  | Se requieren tres tipos de reactivo: espumantes, colectores y controladores          |
| **dentro de** (1) inside +NP       | dentro del molde                                                                     |
| **PATIENT**                       |                                                                                        |
| **de** +SN (11) of +NP             | de unas materias primas                                                             |
| **en** +SN (6)                    | en una pasta o barbotina                                                            |
| **in** + NP                       |                                                                                        |
| **a** +SN (5)                     | al material                                                                          |
| **to** +NP                        |                                                                                        |
| **se aplica/n a** (1) it is applied| se aplica en el esmalte                                                              |
| **MATERIAL STATE**                |                                                                                        |
| **en estado** (1) in X state       | en estado plástico                                                                  |
| **en forma de** (1) in form of     | en forma de barbotina.                                                               |
| **en** +SN (1)                    | en polvo                                                                             |
Table 4. Recurrent syntactical structures in the features of ceramic processes

| Pattern            | Examples from corpora                                      |
|--------------------|-----------------------------------------------------------|
| **OBJECTIVE**      |                                                           |
| infinitive clause  | aplicar un color sobre una superficie                     |
| (25)               |                                                           |
| **PROCEDURE**      |                                                           |
| simple clause      | este se sumerge en la barbotina y se escurre               |
| (23)               |                                                           |
| infinitive         | someter un material a la acción del calor a temperaturas altas. |
| (11)               |                                                           |
| gerund             | añadiéndole componente blanco o agente blanqueante.       |
| (10)               |                                                           |
| NP                 | colocación de piezas unas sobre otras                     |
| (3)                |                                                           |
| **MATERIAL STATE** |                                                           |
| adjective          | bizcochado                                                 |
| (10)               |                                                           |

5. Discussion

We analyzed 222 terminographic definitions belonging to the conceptual groups: ceramic defects and ceramic processes. In these definitions, we detected 13 types of conceptual features: OBJECTIVE, PROCEDURE, PATIENT, INSTRUMENTS, PREVIOUS PHASE, NEXT PHASE, MATERIAL STATE, PHYSICAL ASPECT, ZONE, CAUSE, PHASE, METHOD and PRODUCT. We identified some patterns which have been used to introduce the features into the definitions. The results of this analysis show problems relating to the variety of the feature expression in the definitions, absence of a common genus in the definitions of the same group and other problems which previous studies identified [15, 8], such as the polysemy of patterns, morphological variants and the non-contiguity of the elements of the pattern. We now go on to explain these problems in the following sections.

- Diversity in the linguistic expression of the features in the definitions

The heterogeneity of these definitions affects both the conceptual and linguistic levels. At the conceptual level, we can see that two ceramic defects are not defined when the same features in the definitions are used. For example, in the definition of the defect dunting, the features PHYSICAL ASPECT, CAUSE and STAGE are described. However in the definition of the defect conicity, only the feature PHYSICAL ASPECT is used to describe this concept. At the linguistic level, we can find a wide range of linguistic markers and structures to express the same kind of feature. For example, the feature PROCEDURE is expressed in 10 definitions of ceramic processes by means of a gerund clause. However, in 23 cases, this feature is a sentence which does not show any common pattern. The only regularity found in these cases is the semantic

proximity of the verbs used to describe this feature which are mainly verbs of action (to immerse, to submerge, to project, to pulverize, to attract, etc.), as seen in the following examples: 1. este se sumerge en la barbotina y se escurre; 2. el esmalte se proyecta sobre la pieza cerámica; 3. el producto bizcochado se sumerge en una suspensión de los ingredientes del esmalte en agua; 4. las partículas que han de ser pulverizadas se les da una carga electrostática opuesta a la de la pieza a esmaltar; esta atrae a las partículas hacia la pieza. Semantic annotation would be necessary to automatically extract this feature from the definitions.

- Absence of a genus in the definitions

The formal structure of the analyzed definitions does not always respect the established model of genus and differentia. Many definitions do not include the genus or hypernym of the conceptual group. The first position is filled with a specific feature which cannot be preceded by a linguistic marker.

For example, around 70% of the definitions of ceramic defects do not have the genus defect or imperfection, but include the feature PHYSICAL ASPECT in the first position which is expressed by means of a noun phrase without a linguistic marker, as shown in the following definitions:

rebaba. Delgada cresta (Eng. ‘thin crest’) [PHYSICAL ASPECT] en la superficie o en el borde de un objeto, formada generalmente por la penetración de vidrio entre las partes del molde.

desventado. Grieta (Eng. ‘crack’) [PHYSICAL ASPECT] en la pieza cocida por enfriamiento demasiado rápido.

- Polysemy of linguistic markers

In certain cases, a linguistic marker introduces a different type of feature into the definitions. For example, the structure por + noun phrase is used to express the feature CAUSE, as in por enfriamiento demasiado (Eng., ‘due to fast cooling’), but the PROCEDURE is also used, as in por inmersión de la pieza (Eng., ‘by immersing the tile’). The same happens with the linguistic marker consistente en (Eng., ‘consisting in’) which precedes the feature PHYSICAL ASPECT of a defect (consistente en una textura punteada (Eng. ‘consisting in a dotted texture’)), as well as the OBJECTIVE of a process (consistente en recubrir el acero o hierro con cinc (Eng. ‘consisting in covering the steel or iron with zinc’)).

- Morphosyntactic variety of linguistic markers

Many patterns do not show a unique form in the definitions because they can be expressed in any of their morphological variants. This is especially common in the verbal patterns as they appear to be conjugated in different forms, as the example shows.
lo que origina una estructura vesicular en el material que se calienta
puede originarse en el mismo esmalte, originadas por impurezas,
• Non contiguity of the elements of the marker

We have also found some cases in which the marker is interrupted by another element. For example, the marker formad* por appears in the text with an adverb between the verb and the preposition (formada generalmente por la penetración del vidrio... (Eng., ‘formed generally by the penetration of the glass’))

Results show that the definitions of these dictionaries lack uniformity and that some of the patterns identified in the expression of the features offer some problems that complicate the automatic extraction of information.

6. Conclusions
In this paper we describe the linguistic expression of the different features (such as PHYSICAL ASPECT or CAUSE, included in the definitions of ceramic dictionaries. The aim of the study was to obtain a set of linguistic markers or syntactic patterns for these features. Results reveal a huge heterogeneity in the linguistic realization of the conceptual features in the definitions and the patterns identified show problems such as polysemy and morphological variety.

The ONTODIC project aims to design a tool for computer-assisted terminography which will help the terminographer to create definitions of specialized concepts. These definitions will be based on a template which will guarantee consistency and explicitness. One of the template elements is a restricted set of linguistic markers which will introduce each type of feature into the definitions. This analysis has helped us to observe the naturally occurring markers in the definitions. In future works, we will select one or more linguistic pattern(s) for the expression of each feature in the differentia in our tool for computer-assisted terminography.

7. References
[1] A. Alcina. Metodología y tecnologías para la elaboración de diccionarios terminológicos onomasiológicos, in A. Alcina, E. Valero y E. Rambla (eds.): Terminología y sociedad del conocimiento, Peter Lang, Bern, 2009.
[2] A. Alcina y E. Valero. Análisis de las definiciones del diccionario cerámico científico-práctico. Sugerencias para la elaboración de patrones de definición, Debate Terminológico, 4, 2008.
[3] A. Alcina, V. Soler and J. Granell. Translation technology skills acquisition. Perspectives, Studies in Translatology, 15 (4), 2007.
[4] H. Alshawi. Analysing the dictionary definitions, in B. Boguraev and T. Briscoe (eds.), 1989.
[5] R.A. Amsler. The structure of the Merriam-Webster pocket dictionary, The University of Texas, 1980.
[6] G. Barnbrook. Defining Language. A local grammar of definition sentences, John Benjamins, Amsterdam, 2002.
[7] C. Barrière. From a Children's First Dictionary to a Lexical Knowledge Base of Conceptual Graphs, Ph.D. Thesis, Ottawa University, 1997.
[8] C. Barrière. Investigating the Causal Relation in Informative Texts, Terminology, 7(2), 2002.
[9] B. Boguraev and T. Briscoe (eds.). Computational Lexicography for Natural Language Processing. Longman, London, 1989.
[10] N. Calzolari. Acquiring and representing semantic information in a Lexical Knowledge Base. Proceedings of the ACL SIGLEX Workshop on Lexical Semantics and Knowledge Representation, Springer-Verlag, 1992.
[11] N. Calzolari. Detecting patterns in a Lexical Data Base. Proceedings of the 10th International Conference on Computational Linguistics and 22nd annual meeting on Association for Computational Linguistics, Stanford, 1984.
[12] J. F Chiti. Diccionario de cerámica. Buenos Aires, Condorhuasi, 1984.
[13] C. Guillel and M. C. Guillel. Diccionario cerámico científico-práctico, Castellón, Sociedad española de cerámica y vidrio, 1987.
[14] V. Malaisé, P. Zweigenbaum, B. Bachimont. Mining defining contexts to help structuring differential ontologies, Terminology, 11 (1), 21–53, 2005.
[15] E. Marshman, T. Morgan and I. Meyer. French patterns for expressing concept relations, Terminology, 8 (1), 2002.
[16] I. Meyer, E. Karen and D. Skuce. Systematic concept Analysis within a Knowledge-Based Approach to Terminology, in S. E. Wright and G. Budin (eds.). Handbook of Terminology Management, John Benjamins, Philadelphia, 98-118, 1997.
[17] J. Pearson. Terms in Context. John Benjamins, Amsterdam, 1998.
[18] J. Sager and M.C. L’Homme. A model for the definition of concepts: rules for analytical definitions in terminological databases, Terminology, 1(2), 351- 373, 1994.
[19] G. Sierra, R. Alarcón, C. Aguilar and C. Bach. Definitional verbal patterns for semantic relation extraction, Terminology 14(1), 74-98, 2008.
[20] G. Sierra and J. McNaught. Design of an onomasiological search system: A concept-oriented tool for terminology, Terminology, 6(1), 1-34, 2000.
[21] Sociedad española de cerámica y vidrio. Terminología de los defectos del vidrio. Madrid, 1973.
[22] V. Soler and A. Alcina. Patrones léxicos para la extracción de conceptos vinculados por la relación parte-todo en español, Terminology, 14(1), 2008.
[23] Y. Wilks, D. Fass, C. Guo, J. McDonald, T. Plate and B. Slator. A tractable machine dictionary as a resource for computational semantics, in B. Boguraev and T. Briscoe (eds.), 1989.

