LESS Can Indeed Be More: Linguistic and Conceptual Challenges in the Age of Interoperability
Sara Carvalho, Rute Costa, Christophe Roche

To cite this version:
Sara Carvalho, Rute Costa, Christophe Roche. LESS Can Indeed Be More: Linguistic and Conceptual Challenges in the Age of Interoperability. 12th International conference on Terminology and Knowledge Engineering, Jun 2016, Copenhagen, Denmark. 12th International conference on Terminology and Knowledge Engineering, 2016, 12th International conference on Terminology and Knowledge Engineering. <hal-01354331>

HAL Id: hal-01354331
https://hal.archives-ouvertes.fr/hal-01354331
Submitted on 18 Aug 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
LESS Can Indeed Be More:
Linguistic and Conceptual Challenges in the Age of Interoperability

Sara Carvalho\textsuperscript{123}, Rute Costa\textsuperscript{23}, Christophe Roche\textsuperscript{32}

\textsuperscript{1} School of Technology and Management – University of Aveiro
R. Comandante Pinho e Fretas, 28 3750-127 Águeda – Portugal
sara.carvalho@ua.pt

\textsuperscript{2} NOVA CLUNL – Faculty of Social Sciences and Humanities – Univ. NOVA de Lisboa
Av. de Berna, 26-C 1069-061 Lisboa – Portugal
rute.costa@fcsh.unl.pt

\textsuperscript{3} Condillac Research Group – LISTIC – Université de Savoie Mont Blanc
Campus Scientifique 73376 Le Bourget du Lac – France
christophe.roche@univ-savoie.fr

Abstract. The advent of the Semantic Web and, more recently, of the Linked Data initiative, has paved the way for new perspectives and opportunities in Terminology, namely regarding the operationalization of terminological products. Within the biomedical domain, changes have been substantial in the past decades and at their heart stand the current challenges regarding the production, use, storage and dissemination of medical data, information, and knowledge. In a context where biomedical terminological resources are becoming increasingly concept-oriented, terminology work should reflect a double dimension (both linguistic and conceptual) that may, in turn, support the aspired operationalization and interoperability in this field. Therefore, the purpose of this paper is to present a case study, based around the concept of \textquotedblleft Laparoendoscopic single-site surgery\textquotedblright, in which a methodology anchored in Terminology’s double dimension aims to contribute to the enrichment of the Systematized Nomenclature of Medicine - Clinical Terms (SNOMED CT).

Keywords. terminology’s double dimension; interoperability; laparoendoscopic single-site surgery; SNOMED CT.

1 Introduction

The advent of the Semantic Web and, more recently, of the Linked Data initiative, has paved the way for new perspectives and opportunities in Terminology, namely in what concerns the operationalization of terminological products. The increasingly collaborative work involving Terminology and ontologies – in the sense of Knowledge Engineering (KE) – has led to the development of numerous resources in several areas of knowledge, one of them being Medicine.
Within the biomedical domain, changes have been substantial in the past decades: on the one hand, health care provision has become more technology-based, with computerized examinations, procedures, prescriptions and health records. Furthermore, growing digital literacy has brought the patients into the driver’s seat, where they have been playing a more active – and empowered – role. On the other hand, ageing population and the dramatic decline of the old-age support ratio have contributed to more pressure on public health expenditure, leading to the existing debate around the sustainability of social security systems and their role in health care. At the heart of all these issues stand the current challenges regarding the production, use, storage and dissemination of medical data, information, and knowledge. The ability to provide secure, reliable, efficient and cost-effective ways to process and exchange clinical information among all the stakeholders has emerged as the cornerstone of eHealth initiatives worldwide, with interoperability as one of the key elements\(^1\).

In a context where biomedical terminological resources are becoming increasingly concept-oriented, it is of paramount importance for terminology work to reflect a double dimension (both linguistic and conceptual) that may support the aspired operationalization and interoperability in this field. It is believed that Terminology’s input in the representation, organization, dissemination and, therefore, in the stabilization of specialized knowledge should be taken into account.

Hence, the purpose of this paper is to present a case study, based around the concept of \(<\text{Laparoscopic single-site surgery}>^2\), in which a methodology anchored in Terminology’s double dimension aims to contribute to the enrichment of a particular biomedical terminological resource: the Systematized Nomenclature of Medicine – Clinical Terms (SNOMED CT). The concept under analysis integrates the wider scope of the EndoTerm project, presented in previous papers\(^3\). This article will thus be structured as follows: Section 2 will provide a brief overview of the theoretical background to this case study; Section 3 will be dedicated to SNOMED CT, particularly its logical and concept models; Section 4 will focus on the case study around the concept of \(<\text{Laparoendoscopic single-site surgery}>^4\), followed by some concluding remarks.

## 2 Two sides of the same coin: Terminology’s double dimension

As mentioned above, the double dimension approach, which comprises both a linguistic and a conceptual dimension that are interrelated, has been described by Roche et al. (2009), Roche (2012, 2015), Costa (2013), and by Santos and Costa (2015). According

---

2. A type of surgical procedure that is becoming more and more prevalent in several medical specialties. It is also known as LESS surgery.
3. Cf. Carvalho, Roche, and Costa (forthcoming); Carvalho, Roche, and Costa (2015).
4. Throughout this paper, concepts will be capitalized and written between single chevrons, whereas terms will be presented in lower case and between double quotation marks (cf. Roche 2015).
to Roche (2015: 136), Terminology is “both a science of objects and a science of terms”. For Costa (2013), it is precisely this double dimension, as well as the study of the relationship between one and the other, that grants Terminology its place as an autonomous scientific subject.

This double dimension perspective implies, therefore, that both the experts’ conceptualizations of a given subject field and the discourses produced by them must be taken into account in terminology work, thus leading to a complementarity of two fundamentally different dimensions. Consequently, both specialized texts and expert collaboration constitute invaluable resources in terminological work, provided that there is a supporting theoretical and methodological framework that allows the terminologist to maximize the potential within each dimension, and mostly of the synergies resulting from their interaction.

3 Current biomedical terminological resources and interoperability: the example of SNOMED CT

Medicine is currently undergoing significant challenges regarding the way clinical information and knowledge are produced, used, stored and shared. In recent years, many biomedical terminological resources have been designed or redesigned in order to incorporate ontology-based elements, thus evolving from “simple code-name-hierarchy arrangements, into rich, knowledge-based ontologies of medical concepts” (Cimino 2001). Concept-orientation has become one of the key principles of current biomedical resources and was, in fact, one of the twelve desiderata that, according to Cimino (1998), should support biomedical terminological systems in the 21st century.

One of these resources is SNOMED CT, currently owned and distributed by the International Health Terminology Standards Development Organization (IHTSDO). It is a comprehensive, multilingual health care terminology that due to its description-logic basis, supports the representation of clinical content in electronic health formats (namely Electronic Health Records – EHRs) in a consistent, reliable and computer-readable way.

This resource has been built around three main components: the concepts, which represent clinical meanings, are organized into hierarchies, ranging from general to specific; the descriptions, which provide the human readable form of a concept, comprise the Fully Specified Name (FSN), representing “a unique, unambiguous description of a concept’s meaning”\(^6\), and the synonym (SYN). Each concept may have multiple synonyms, but only one is marked as “preferred” in a given language, whereas the remaining synonyms are marked as “acceptable”; finally, the relationships, which connect concepts to other related concepts, are used to logically define the meaning of a concept in a computer-processable way. There are two main types of relationships: subtype, or is_a relationships, that form the basis of SNOMED CT’s hierarchies, and attribute relationships, that associate the source concept (e.g. abscess of heart) with the value

\(^5\) For more information, see http://www.ihtsdo.org/snomed-ct/what-is-snomed-ct

of a defining characteristic. The characteristic (attribute) is specified by the relationship itself (e.g. |finding site|) and the value is provided by the destination concept (e.g. |heart structure|). Each one of these three components (concepts, descriptions and relationships) has its own unique numeric identifier.

SNOMED CT has been selected as the basis of our proposal for several reasons: firstly, the concept-based nature of this resource and the fact that it aims to integrate the linguistic and the conceptual, while preserving their fundamental differences, appear to be consistent with Terminology’s double dimension and its core principles; secondly, SNOMED CT’s structure allows post-coordination, i.e. more complex concepts may be created from a set of more primitive components. In this type of system, also called compositional (cf. Coiera 2015; Duclos et al. 2014), there is no need to create all the elements in advance, but rather to ensure that all the basic building blocks exist. It is therefore possible to represent a given clinical content even when the precise concept is not present in SNOMED CT. This representation may occur via a standard compositional grammar that is both human-readable and computer-processable, thus enabling interoperability. Moreover, this compositional approach to concept representation requires the definition of a set of logical rules (constraints) that will govern the way concepts and relationships can be combined, in order to prevent nonsense representations; thirdly, and unlike other resources of its kind, SNOMED CT is not limited to hierarchical concept relations; finally, the concept under analysis does not exist in this terminological system, so it is believed this proposal could contribute to enrich SNOMED CT’s content.

4 LESS: a Brave New World for surgery?

As stated above, the development of the EndoTerm project, which was described in depth in Carvalho, Roche, and Costa (2015) and which aims at the creation of a multilingual terminological resource based around the concept of <Endometriosis>, led to the study of single port surgery, a relatively recent type of minimally invasive surgery. The further analysis of the concept pointed towards a lack of terminological consensus among the expert community, with a plethora of terms coined by individual groups and organizations. In fact, more than 20 have been identified in the literature.

---

7 As well as with Ontoterminology (cf. Roche et al. (2009); Roche (2012); Roche (2015); Carvalho, Roche, and Costa (2015).
8 A concrete example of this compositional grammar will be presented in the next section.
9 Also known as the categorial structure (cf. ISO 17115: 2007). For example, has_site should occur only between concepts related to morphology and concepts referring to topography (e.g. pyelonephritis is_a infection (morphology: -itis) which “has_site” kidney (topography: – nephr-).
10 Due to space constraints, it was not possible to include a table with all the collected designations in this paper (a total of 22). However, they can be found in Box et al. (2008); Gill et al. (2010); Autorino et al. (2011); Ramesh, Vidyashankar, and Dimri (2014); Georgiou et al. (2012); Springborg and Fader (2015); Escobar and Falcone (2014), just to name a few.
In order to solve this terminological dispersion, in 2008 a multidisciplinary medical consortium decided to standardize the terminology in the field and proposed the term “laparoendoscopic single-site surgery” (also known as “LESS surgery”) as the one that most accurately depicted this surgical procedure. In addition, LESSCAR’s White Paper highlighted the characteristics that the concept should encompass: 1) a single entry port (or incision); 2) applicability to multiple locations (abdominal, thoracic or pelvic); 3) umbilical or extramumbilical access; 4) type of surgery (laparoscopic, endoscopic, or robotic); 5) type of surgical approach (percutaneous intraluminal and percutaneous transluminal). The group also required that all scientific publications on LESS surgery should include a “mandatory descriptive second line”, with details about the number and type of ports used, the type of laparoscope used, and the type of instruments used. (Gill et al. 2010).

Within the scope of the EndoTerm project, the gathered data concerning LESS surgery provided sufficient ‘food for thought’ in order to constitute a terminological case study. A number of questions interconnecting the linguistic and conceptual dimensions arose from the data analysis: i) are all the gathered terms actual synonyms from a terminological standpoint, i.e. representing the same concept and being interchangeable in all contexts (cf. ISO 1087-1, 2000)? ii) what usage has the expert community in the field of Gynecology been making of these designations? iii) knowing that texts do not contain concepts themselves, but the linguistic usages of the terms that designate them, what type of information could be extracted from a set of natural language definitions? And in what way would that match a concept map validated by subject field experts? iv) given that the concept of <Laparoendoscopic single-site surgery> does not exist in SNOMED CT, what additional information would be necessary for its inclusion in this resource and how could it be represented in a way that enables interoperability?

In the literature, the designations used to refer to LESS surgery are often depicted as synonyms, which, from a terminological point of view, as mentioned earlier, raises the dilemma of whether apples are indeed being compared to apples. An analysis of the terms shows that the notion of a single access to the body (“single incision”, “single access”, “single port” or “single site”) seems to play a central role in this type of surgery. Two of the terms, however, refer to “incisionless” and to “natural orifice” surgery, respectively, which indicates the inexistence of an external incision, hence opposing the notion that prevails in the remaining designations. Additional information is provided by most terms as regards the location of the incision (“umbilical” or the more specific “transumbilical”), the type of surgery (the more generic “minimally invasive surgery” or the more specific “endoscopic” and, going further down the hierarchy, “laparoscopic”), and the use of a given type of equipment (“video”, “conventional equipment-utilizing”). This seems to point towards the idea that not all these designations are in fact representing the same concept, but a more thorough analysis, which is not within

---

11 The Laparoendoscopic Single-Site Surgery Consortium for Assessment and Research (LESSCAR), that published a consensus statement with the main conclusions of that meeting (Gill et al., 2010). The Urologic NOTES Group also endorsed LESS surgery as the designation for single-port surgery (Box et al., 2008).

12 The medical specialty more actively devoted to the diagnosis and treatment of Endometriosis.
the scope of this article, would be necessary in order to confirm this hypothesis and further develop it.

In order to get a glimpse of the actual usage of these terms among the community of subject field experts, a search was conducted in MEDLINE/PubMed® with the full forms of the 22 collected designations and resorting to the following search expression: ("Term" [All Fields] AND gynecology [All Fields]) AND ("2010/01/01" [Date-Publication]: "2016/03/01" [Date-Publication]). The aim was to see which terms have been more widely used in scientific, peer-reviewed papers within Gynecology since 2010 (the date of publication of LESSCAR’s White Paper). The results showed that “laparoendoscopic single-site surgery” has become the most commonly used designation (with 90 results), followed by “single-port access” (45), “single-port laparoscopy” (29), “single-incision laparoscopic surgery” (27), and “single-port laparoscopic surgery” and “single-port surgery” (23 results each).

From the 90 scientific articles for “laparoendoscopic single-site surgery”, only 15 had the full text freely available and were in English, so they constituted the selected corpus. The AntConc corpus analysis tool allowed the analysis of a set of definitions of “laparoendoscopic single-site surgery” and the subsequent extraction of data, leading to the following lexical network:

![Lexical network created with data extracted from the corpus](image)

In this particular example, it seems that resorting only to a corpus-based approach, though useful, might be insufficient to fully grasp the notion of LESS surgery. Thus, there was a search for additional information in a set of biomedical terminological resources. Some of the resources were entirely hierarchical (such as ICD-10 and MeSH), others contained non-hierarchical relations as well (e.g. SNOMED CT and UMLS). Current procedure classifications were also consulted, namely the NOMESCO Classification of Surgical Procedures (2012); the German Procedure Classification (Operationen- und Prozeduren schlüssel – OPS, 2016 version); the OPCS Classification of Interventions and Procedures, version 4, used by the UK’s National Health Service; and the French Classification Commune des Actes Médicaux (2016 update). Being a type of surgery that is estimated to account for 50-80% of current surgeries in some medical specialties (particularly urology and gynecology) (cf. Gill et al., 2010), it was surprising to realize that the concept as such does not exist at the moment in any of the

---

13 Available at [http://www.laurenceanthony.net/software.html](http://www.laurenceanthony.net/software.html).

14 Is a and is performed through were actually present in the corpus and hence express lexical relations. The former should therefore not be confused with the conceptual is_a relation.
consulted resources. Although data on <Laparoscopy> or <Minimally invasive surgery> are available, there is nothing that refers specifically to a single incision, which would allow concept differentiation. The inclusion of additional content about <laparoendoscopic single-site surgery> in one of these resources was believed to be pertinent, and SNOMED CT has been chosen based on the arguments presented earlier.

Bearing all of this in mind, supplementary searches were conducted, showing the need to go beyond the verbal and incorporate non-verbal (images or diagrams) as well as multimodal elements about LESS surgery (within the context of Gynecology). The data analysis led to the creation of a set of concept maps using the OTE tool©, which were then validated by two senior expert gynecologists. Due to space limitations, only one of the maps will be shown. The map below (Fig. 2) depicts the concept under analysis and aims to position it within the broader concept of <Surgical procedure> by making use of a specific differentiation, Aristotelian-based approach. It confirms that the lexical network from Fig. 1, although incomplete, contains elements that may indeed correspond to relevant characteristics of the concept under analysis. As a matter of fact, the existence of a single skin incision constitutes the essential characteristic (cf. ISO 1087-1: 2000) of this type of surgical procedure.

Fig. 2. Concept map of <Laparoendoscopic single-site surgery>

Taking into account both the linguistic and conceptual information gathered thus far, and after analyzing SNOMED CT’s categorial structure for procedure concepts, it is

---

15 Namely medical video articles, a new type of scholarly communication that has been more thoroughly described in Carvalho, Roche, and Costa (forthcoming).

16 Created by the Condillac research group, from Université de Savoie Mont Blanc. The maps, as well as the tool, have been described in more detail in Carvalho, Roche, and Costa (2015).
believed that this resource could benefit from the inclusion of the concept <Laparoendoscopic single-site surgery>, given the increasing prevalence of this type of surgery in some medical specialties and the likely need to refer specifically to LESS surgery in EHRs. Therefore, our proposal would be as follows: as regards the descriptions (the linguistic dimension), the choice of the Fully Specified Name and respective Synonyms would respect the position issued by LESSCAR and supported by our MEDLINE/PubMed® searches, as seen below (Fig. 3).

As for the conceptual dimension, which is the basis of concept definitions in SNOMED CT, the biggest challenge lies in the current absence, in this resource, of any attribute-value relationship referring to the essential characteristic of the concept in question, i.e. the single skin incision. Although the concept of <Incision of skin (procedure)> exists, its subtype concepts are related to the location of the incision, with no data concerning the number of incisions. The same happens with <Incision – action (qualifier value)>, which refers to the method used in practically all surgical procedures. Since <Incision of skin (procedure)> would not work as a valid destination concept for the attribute relationships used to define procedure concepts, as it would conflict with the domain constraints, it is believed that the concept of <Single incision – action (qualifier value)> should be introduced in SNOMED CT, in order to enable concept differentiation and, hence, concept definition.

The concept of <Laparoendoscopic single-site surgery> could therefore be defined through a combination of is_a and attribute relationships, represented in both a human- and computer-readable way via SNOMED CT’s compositional grammar, which supports interoperability. The following proposal (Fig. 4) resorts, as much as possible, to

---

17 For consistency purposes, SNOMED CT’s terminology will be maintained in our proposal. Therefore, instead of adopting the notion of designation (cf. ISO 1087-1), the original expression “description” will be used.

18 For further information, cf. IHTSDO (2015).
currently existing concepts, descriptions and relationships. The suggestions have been signaled in red.

![Fig. 4. <Laparoendoscopic single-site surgery> using SNOMED CT’s compositional grammar](image)

The first three concepts represent the is_a relationships (types of surgery) and are followed by a refinement, which is introduced by a colon and consists of a sequence of one or more attribute-value pairs. The attribute is separated of the value by an equals sign and if there is more than one value for the same attribute, the plus sign is added. The different attribute-value pairs are separated by commas. Curly braces represent grouping of attributes within a refinement, for example to indicate that a given method applies to a specific site.

5 Concluding remarks

By resorting to a case study, this paper aimed to reflect upon the fact that analyzing the conceptualization of a given subject field and the corresponding discourses produced by the expert community may result in representations that do not always match, but both play a vital role in terminology work: through ontologies, conceptualization proposals open new possibilities in terms of interoperability by resorting to the Semantic Web and W3C standards; albeit with vagueness and inconsistencies, the discourses provide fundamental access to the expert community as a way to stabilize knowledge in different areas of expertise, which is particularly relevant in new techniques or approaches, as is the case of LESS surgery. When anchored in this double dimension, terminology work may contribute to further enhance that stability and, consequently, the quality of specialized communication.
Acknowledgements

This research has been 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/03213/2013.

References


