JoSTrans
The Journal of Specialised Translation

www.jostrans.org · ISSN: 1740-367X

Wermuth, M.-C., J. Walravens, M. & H. Lambot, M.-A. (2022). Collaboration and communities of practice in the field of medical ontology translation. *The Journal of Specialised Translation, 37*, 75-98. https://doi.org/10.26034/cm.jostrans.2022.105

Collaboration and communities of practice in the field of medical ontology translation

Maria-Cornelia Wermuth, KU Leuven
Michel J. Walravens, MD
Marie-Alexandra H. Lambot, CHU Saint-Pierre

ABSTRACT

This article provides insight into the collaborative translation of medical ontologies and how translators' needs, commitment and desire to learn and share knowledge can be the seeds of a community of practice to support this type of translation. It is based on the experience of a project to translate 7,500 respiratory system concepts from SNOMED CT® into French. The project was guided by the following principles: (1) translators are bilingual medical experts in the field in which the terms to be translated are used; (2) linguistic advice is available during translation; (3) translating experts accept, correct or replace terms already translated by a web-based translation tool using a prepared lexicon. The project results were characterised by low acceptance of computer-assisted pre-translations due to inadequate lexicon preparation and insufficient alignment of the concepts to be translated with the expertise of the translating expert. These issues were solved to a considerable extent by web-based communication between translators, suggesting the need for well-structured collaboration between highly specialised field experts. Based on this observation, we discuss how a community of practice built on the motivation and needs of the translating experts could significantly support the quality and efficiency of medical ontology translation.

KEYWORDS

Collaborative medical translation, communities of practice, ontology-based medical translation, SNOMED CT®, medical concept translation, translation training.

1. Introduction

The Systematized Nomenclature of Medicine – Clinical Terms (SNOMED CT®) is an English-language clinical reference terminology used worldwide (see https://www.snomed.org/) primarily for recording clinical data, such as diseases, clinical findings and procedures in electronic health records (EHRs) during patient care (Benson and Grieve 2016). SNOMED Clinical Terms® has been used for this article with permission from SNOMED International. All rights reserved. SNOMED CT® was originally developed by the College of American Pathologists. "SNOMED", "SNOMED CT" and "SNOMED Clinical Terms" are registered trademarks of SNOMED International.

The use of SNOMED CT in non-English speaking countries requires translations in different languages. This article presents retrospective analysis of a translation project launched by the French group of the Belgian Community for Support on Clinical Terminologies (CSCT) in response to a call for tender from the National Release Centre (NRC), responsible for the translation and release of SNOMED CT in Belgium. (For more detail on NRCs,

see Bhattacharyya [2016].) The project consisted in providing descriptions of about 7,500 SNOMED CT concepts from the field of the respiratory system in Belgian French (RS project) and was carried out collaboratively by a group of paid volunteers. All of them were domain experts from different medical specialties, and there were no professional translators or linguists involved. The translating domain experts were assigned the role of 'reviewer' since their main task was either to accept, correct or re-translate pre-translated terms. These pre-translated terms were automatically created at the end of the first stage of the project, which involved lexicon building. At the second stage of the project, which involved SNOMED term translation, the terms to be translated were arbitrarily divided into batches. One or more batches were assigned to pairs of reviewers, with reviewer 1 (REV1) doing the first revision of the pre-translated terms and reviewer 2 (REV2) doing the revision of the REV1 work. The work of the reviewers was checked by a team of three supervisors (SUVs) after revision. In some cases, the supervisors also intervened in the course of the revision, either spontaneously or at the request of the reviewers. Reviewers and supervisors thus cooperated to finalise the translations by accepting, correcting or retranslating the pre-translations prepared in stage 1. Although detailed examination of the first stage of the project (lexicon building) is beyond the scope of this paper, it is briefly explained in Section 3.1 because of its importance for the quality and efficiency of the final translations. The focus of this paper is, instead, on the second stage of the project, which involved translation of SNOMED terms.

The group used the Concept-based Medical Term Translation tool (CoMeTT), a web-based translation platform that enables collaboration between translators/reviewers and supervisors. The group formed an organised team led by a project coordinator, who was responsible for practical organisation, and a project manager, who was tasked with instructing and supervising the group. For the purposes of this article, collaborative online translation is therefore used in a broader sense; i.e. as "collaboration between a number of translating and non-translating agents, one or more of which may not be a translator" (Jiménez-Crespo 2017: 18).

The primary goal of CoMeTT is to increase the efficiency of translation quality and speed by presenting translators/reviewers with one or more pre-translated terms for acceptance or correction. Together with the CoMeTT pre-translations, all translations entered by translators/reviewers during this or previous projects or imported from other sources are presented to the reviewers as pre-translations. During the RS project, two important observations could be made: on the one hand, the quality and speed of the (pre-)translations fell short of expectations; and on the other hand, the reviewers, and especially the supervisors, made intensive use of the implemented free text field intended for the exchange of questions and suggestions. The latter observation highlights the participants' need for knowledge sharing and desire for mutual learning during translation and reviewing. This interaction is reminiscent of the notions of "knowing in

action" (Amin and Roberts 2008: 354) and "situated learning" (Lave and Wenger 1991) as described in the literature on communities of practice. It could lead one to assume that the group evolved into a virtual community (Ellis *et al.* 2005) in the course of the project. However, for various reasons to be discussed, this was not the case. Rather, the question is now whether a 'real' community of practice (hereafter, CoP; plural CoPs) should be established to engage with the collaborative group in translation projects in the field of ontology-based medical terminology, thus supporting and optimising the functioning of the group by sharing its knowledge and "lived experience" (Pyrko *et al.* 2017: 405).

The aim of this article is to further explore the added value and practical implementation of such a CoP, taking into account the 'CoP-like' elements that emerged during the collaborative translation process and that are assumed to be essential for the formation of a CoP. For this purpose, we analysed the communication process during the collaboration, including questions from reviewers and supervisors about the appropriateness of term candidates, term definitions, grammar and spelling, and comments on the usability of the user interface. Based on these outcomes, we discuss how, in the context of SNOMED CT translation, a CoP with members from different disciplines (e.g. Medicine, Translation Studies, Terminology Science, or Health Informatics) could be formed to serve different goals, such as *ad hoc* recommendations for translation and teaching.

The article is structured as follows. In Section 2, we introduce SNOMED CT and describe the specific challenges associated with its translation. In Section 3, we describe the translation tool CoMeTT and the steps and distribution of roles during the translation process. In Section 4, we describe the outcomes of the translation process and the group interaction during the translation project. Based on the insights gained through these descriptions, we discuss some possibilities of how a CoP could be put to use in the specific context of translating and teaching (ontology-based) medical terminology. In Section 5, we address the specific characteristics of a web-based collaborative translation group on the one hand, and a CoP for (ontology-based) medical terminology on the other. We further discuss how an integrated collaboration between the two can be achieved, leading to mutual knowledge gain and providing support for the achievement of their respective goals. In Section 6, we summarise our findings.

2. SNOMED CT and concept-based translation

2.1. Formal features of SNOMED CT

SNOMED CT is a reference terminology resource used in electronic health applications (e.g. EHRs, e-prescriptions, laboratory reports) in more than 80 countries, either through national membership of SNOMED International or through affiliate licensees, to enable standardised data exchange between healthcare providers (Spackman and Reynoso 2004; Santos and

Costa 2015; Bhattacharyya 2016) with the aim to improve the quality and safety of patient care. The system consists of three components: (1) concepts, (2) descriptions and (3) relationships. A concept is defined as "a clinical idea to which a unique concept identifier has been assigned" (IHTSDO 2021). Besides these machine-readable unique identifiers, concepts are also represented by human-readable descriptions (single word terms, multiword terms and complex noun phrases) in natural language (IHTSDO 2018; Bhattacharyya 2016).

The concepts and their mutual relationships are represented in the form of a formal ontology, which defines the relationships between the concepts, such as findings, procedures, disorders or events, using OWL Description Logic (Ivanovic and Budimac 2014). There are two types of relationships: hierarchical and attributive. The hierarchical is a relationship relating a (child) concept to a more general (parent) concept. There are also more than 100 attributive relationships representing non-hierarchical meaning aspects such as causative agent, pathological process or finding site. Conceptual relationships are generated in two ways, either directly stated by the human terminologists modelling the new concept or automatically inferred by the SNOMED CT description logic classifier. The representation of the former is referred to as the stated view, and the representation of the latter is referred to as the inferred view of the concept logical description. Since many clinical concepts are inherently multidimensional (Benson and Grieve 2016), concepts can have multiple parents, which creates a polyhierarchical structure (Figure 1).

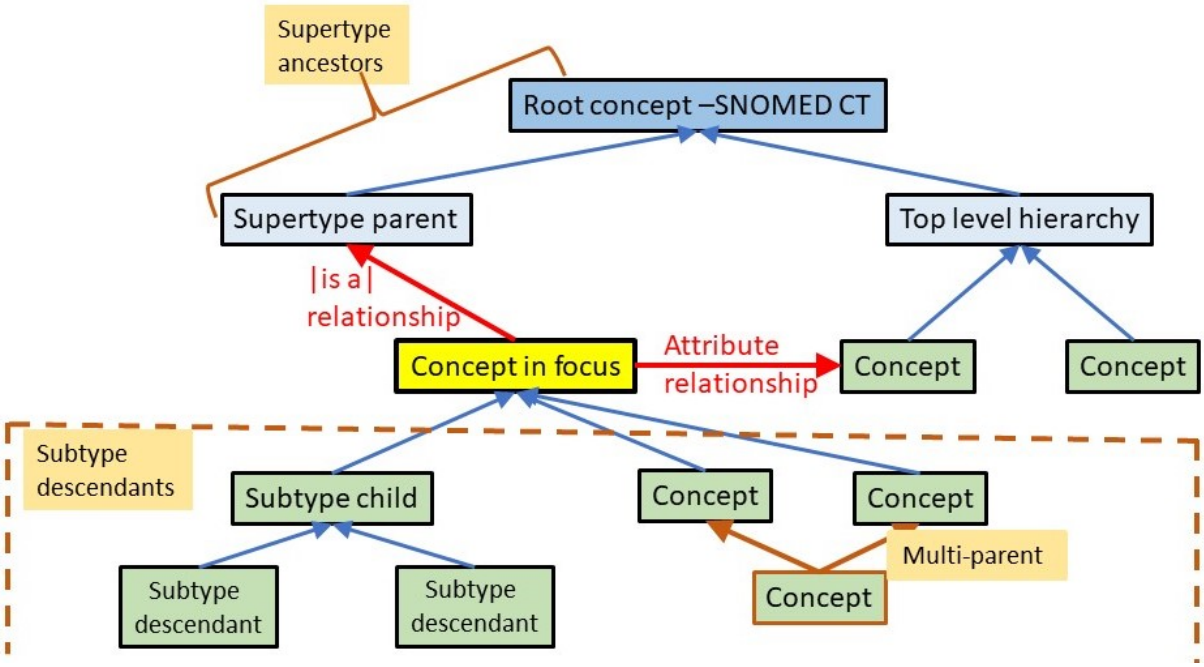

Figure 1. The SNOMED CT hierarchy (IHTSDO 2012a)

Concept descriptions are labels assigned to a SNOMED CT concept. There are three types of descriptions: fully specified name (FSN), synonym and definition. Note that the description terms are written between pipes by convention, and we follow this convention throughout the article.

The FSN is "a description that represents the meaning of a concept in a way that is unambiguous and independent of the context in which it is used" (IHTSDO 2021). An FSN is composed of a term and a semantic tag between parentheses at its end. Take, for example, the FSN |Myocardial biopsy (procedure)|. The tag indicates the (sub)hierarchy, or broad category, to which the concept belongs (e.g. disorder, organism, procedure). This composition makes each FSN unique for each concept, even in cases of FSNs with identical terms that refer to concepts belonging to different hierarchies. For example, |Hematoma (morphologic abnormality)| is the FSN of concept 35566002, that represents what the pathologist sees at dissection, whereas |Hematoma (disorder)| is the FSN of concept 385494008 that indicates the clinical diagnosis of a hematoma on the patient.

A synonym is "a word or phrase that expresses the meaning of a SNOMED CT concept in a specific language" (IHTSDO 2021). Each concept may have multiple synonyms, one of which is marked as preferred to support consistent coding. The preferred term (PT) is "the term deemed to be the most clinically appropriate way of expressing a concept in the specified language context" (IHTSDO 2021). The other synonyms are marked as acceptable. For example, the phrases |Disease caused by 2019 novel coronavirus| and |Disease caused by 2019-nCoV| are acceptable synonyms of the PT |COVID-19|, and all three descriptions refer to the same concept 840539006 |Disease caused by Severe acute respiratory syndrome coronavirus 2 (disorder)|. Figure 2 shows the different SNOMED CT components.

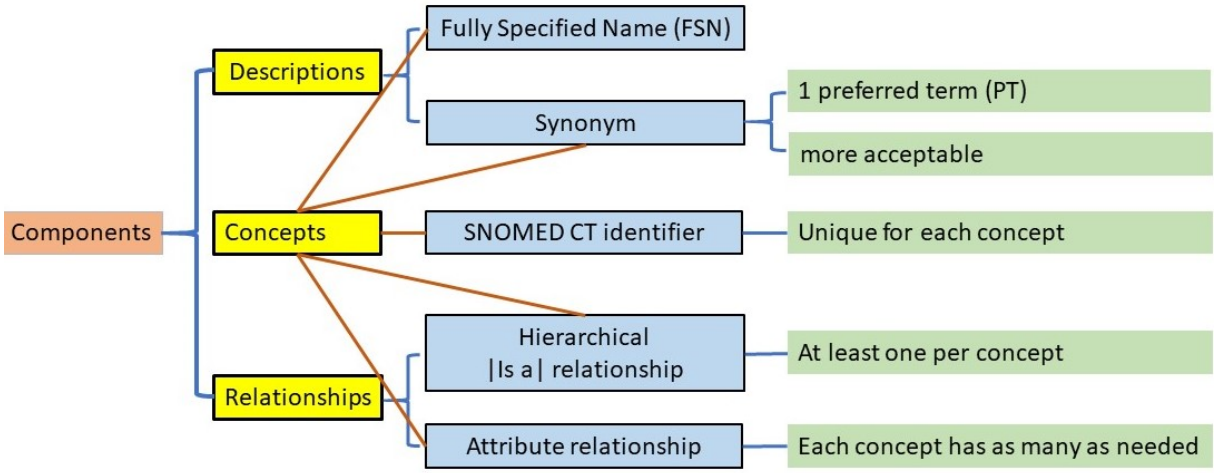

Figure 2. The SNOMED CT logical model (adapted from IHTSDO 2018: Ch. 5)

In the SNOMED CT ontology, the relationships should strive to define a concept sufficiently to distinguish it "from any concepts or expressions that are neither equivalent to, nor subtypes of, the defined concept" (IHTSDO 2021). In this case, the term is fully or sufficiently defined in relation to its immediate supertype(s). A primitive concept is a concept that is not fully or sufficiently defined (IHTSDO 2018: Ch. 3.2). A defining relationship is "a relationship to a target concept that is always necessarily true for any instance of the source concept" (IHTSDO 2021). For example, Figure 3 indicates that the concept 74400008 |appendicitis (disorder)| is sufficiently defined by its four relationships, as each concept to which these defining relationships apply is either the disorder |appendicitis| or a subtype of |appendicitis| (IHTSDO 2018).

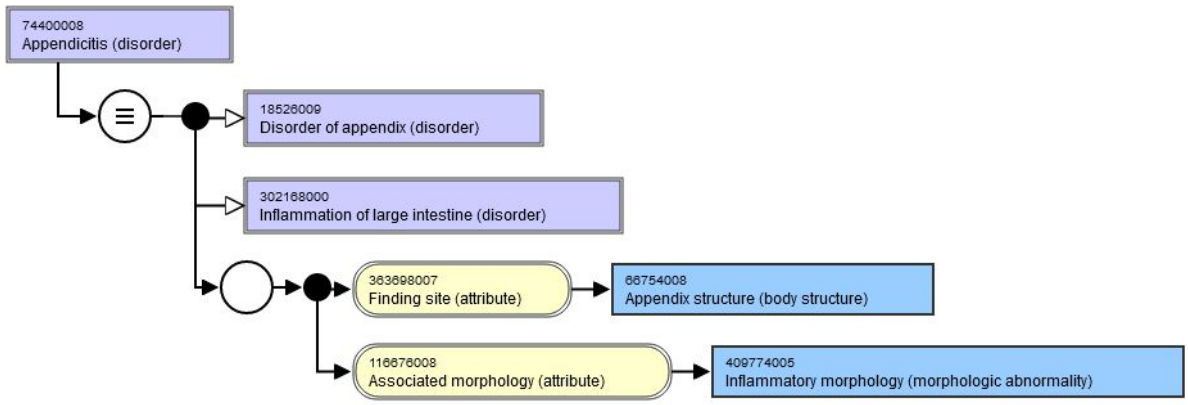

Figure 3. Diagrammatic representation of the relationships, inferred view, of concept 74400008 |Appendicitis (disorder)| as modelled in the International Edition v2021-01-31

2.2. Concept-based translation

High quality translations are of great importance to SNOMED CT, as one of its aims is to contribute to the standardisation of the language for special purposes (LSP) used in medical practice and research. For this reason, international guidelines (IHTSDO 2012a, 2012b) were developed with recommendations for best practice in the management of translation projects. These international guidelines should be used in combination with national guidelines (if available) that take into account language-specific aspects.

Translations should be done by a multidisciplinary team of various actors, such as software developers, health and social care professionals, (medical) translators, reviewers and terminologists, preferably all with knowledge of health informatics and the use of terminology in healthcare, as well as experience of the SNOMED CT ontology. The translation should run through a multi-stage process that is completed with thorough quality checks. Given the highly specialised terminology, recruiting domain experts as translators and/or reviewers is a recommended approach, as their expertise

guarantees validity and clinical acceptance of the translations, which supports the efficiency of the entire translation process.

According to the *Guidelines for Translation of SNOMED CT* (IHTSDO 2012a), translations should be concept-based: before translating a term, the meaning of the underlying concept must be clearly understood. To understand the meaning of a concept, the guidelines recommend the determination of the hierarchical position of the concept in the SNOMED ontology after reading the FSN. However, using the ontology to grasp the concept is problematic for several reasons: (1) Translators who are not medical experts are unlikely to understand the ontological definition, which is an expert-oriented formal representation of meaning; (2) The meaning of a concept may consist of more features than the attributes represented in the ontology, especially for primitive concepts; (3) Each defining attribute is, in turn, a medical concept, the interpretation of which again requires medical expertise, as illustrated by the example in Figure 4.

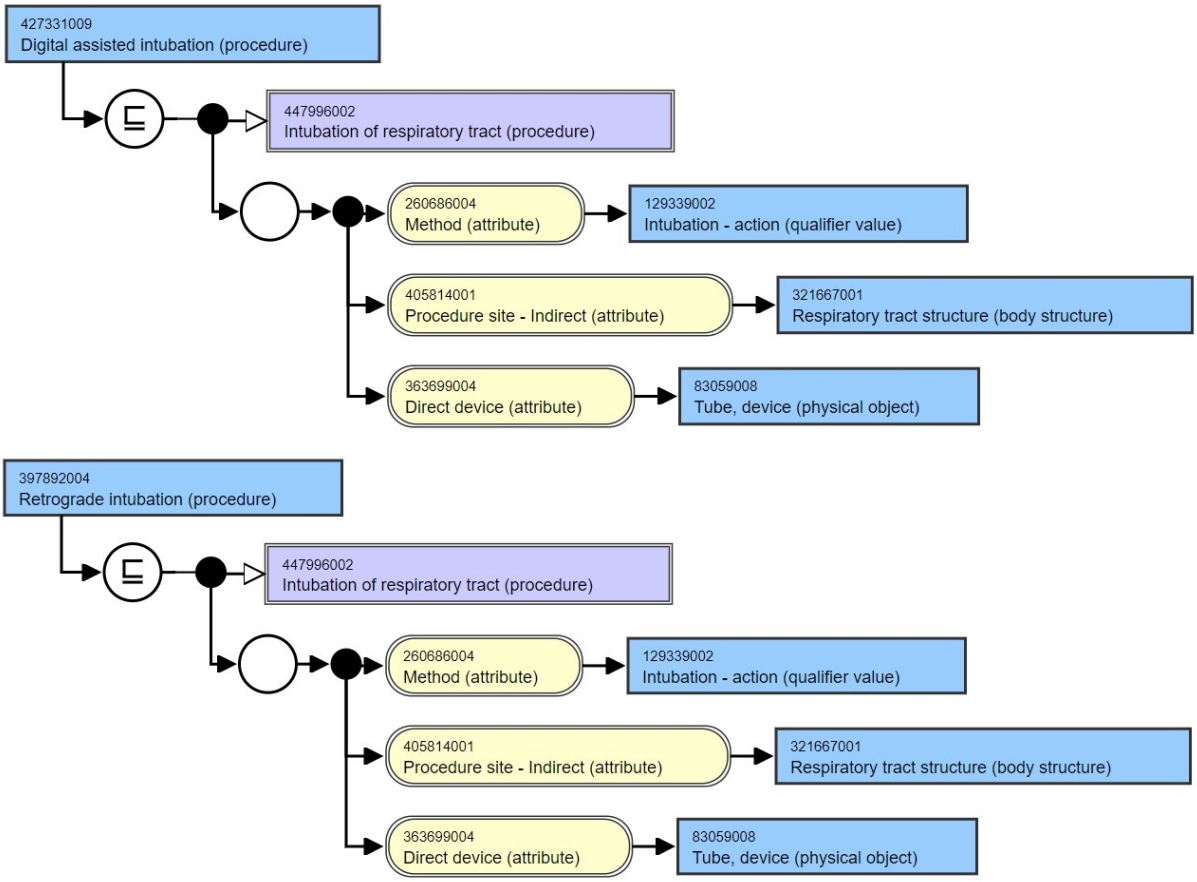

Figure 4. Diagrammatic representation of the relationships, inferred view, of concept 427331009 |Digital assisted intubation (procedure)| and its sibling 397892004 |Retrograde intubation (procedure)| as modelled in the International Edition v2021-01-31

The attribute relationships of the two primitive concepts illustrated in Figure 4 do not allow these two concepts to be distinguished from each other. Moreover, each FSN includes meanings (respectively "digital assisted" and

"retrograde") that are not represented by the defining attributes. Only a medical specialist having performed those procedures would know for sure that the word "digital" in the first concept refers to the fact of using one's finger to help set the tube in place into the respiratory tract and does not refer to the use of a digital electronic device.

Given these specific translation challenges, a basic assumption of the current study is that efficient use of the SNOMED CT ontology as a knowledge source for translation purposes requires that the translators not only have both in-depth domain and ontology knowledge, but are familiar with term usage in clinical practice as well.

3. Concept-based Medical Term Translation tool (CoMeTT)

3.1. Database structure

CoMeTT was originally developed as a local tool to support the translation of SNOMED CT, based on principles derived from the developer's experience of translating 362 concepts in the field of rheumatology, most of them systemic autoimmune disease (AID) concepts. These principles are: (1) Adequate and efficient translation of terminological ontologies requires domain experts who are familiar with the meaning of the concept, the conceptual structures of the domain and the terminological usage in clinical practice in both the source and target languages; (2) A linguistic review should follow the translation process and validate the final translation; (3) The translating domain experts should be supported by machine pre-translations based on a translated lexicon and by an interactive, user-friendly interface with links to the SNOMED CT browser (https://browser.ihtsdotools.org/), which provides access to all SNOMED CT releases available in different languages and displays a variety of information on terms, as well as links to relevant medical and scientific websites; (4) The lexicon should be extracted from the source terms, automatically translated, manually corrected and grammatically edited indicating term type (single or multiword term), gender and context for synonyms, if applicable. On this basis, database tables were created (Figure 5).

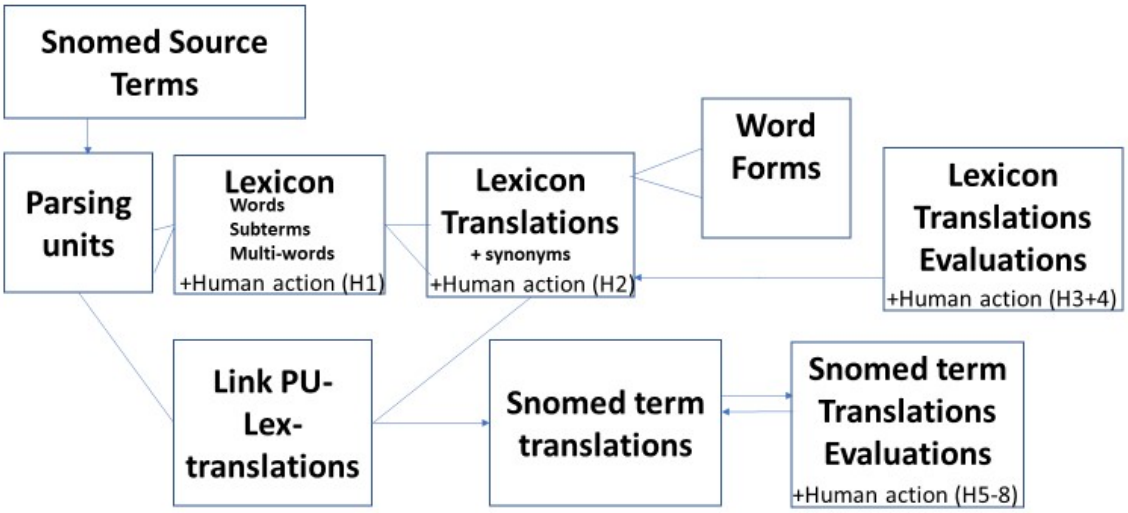

Figure 5. Structure of the CoMeTT database (CDB)

Of relevance to this article, CoMeTT has a built-in collaboration feature that allows reviewers and supervisors to ask questions about the translation of specific concepts. Questions and comments can be entered in a free text field that is automatically linked to the term under discussion. Any response or comment to that term is entered into the same free text field and marked with author, date and time. This collaboration function enables quick feedback on questions and translation problems, supports teamwork and also has a positive effect on group dynamics.

3.2. Efficiency of the lexicon-based approach by domain experts

We briefly summarise here data on the efficiency of the lexicon-based approach within CoMeTT in the context of the full International Edition of SNOMED CT.

The efficiency of lexicon-based translation appears very promising as redundant tasks are avoided since each lexical unit only needs to be translated once. The efficiency of lexicon-based translation is measured by the frequency of a lexical unit in the whole database. Parsing the 336,893 FSNs in the July 2017 International Release yielded 1,383,163 lexical units, of which 68,473 occurred once. The mean frequency of each lexical unit in the totality of the source terms was thus 20.2. The hierarchical tags of the FSNs, one per FSN, have the highest frequencies. There were 113,298 FSNs in the disorder hierarchy of the January 2021 International Release, and thus at least as many disorder tags.

In the RS project on which this article is based, machine translation (Google translate) was used for the translation of single word units such as "ulna"

and a correctness rate of about 75% was achieved. These automatically generated lexicon translations, which were reviewed and corrected by domain specialists, supported the subsequent automatic generation of pre-translations for concepts using these single-word units in their descriptions. If single-word units had more than one meaning, synonyms could be created during the revision of the automatic translation with the support of the translation tool. Part-of-speech tagging of the English source words in the lexicon and gender marking of the translated nouns required only one click per word, resulting in n=10 words per minute. For adjectives, rules for automatic inflection (by gender, singular and plural) were coded in French so that they could be correctly linked to the nouns.

Based on the lexicon, syntactic rules for the automatic translation of terms were created. The screenshot in Figure 6 shows the user interface for lexical translation. For context checking in lexical translation, clicking on a lexicon word displayed the list of all source terms (FSNs) containing the lexicon word.

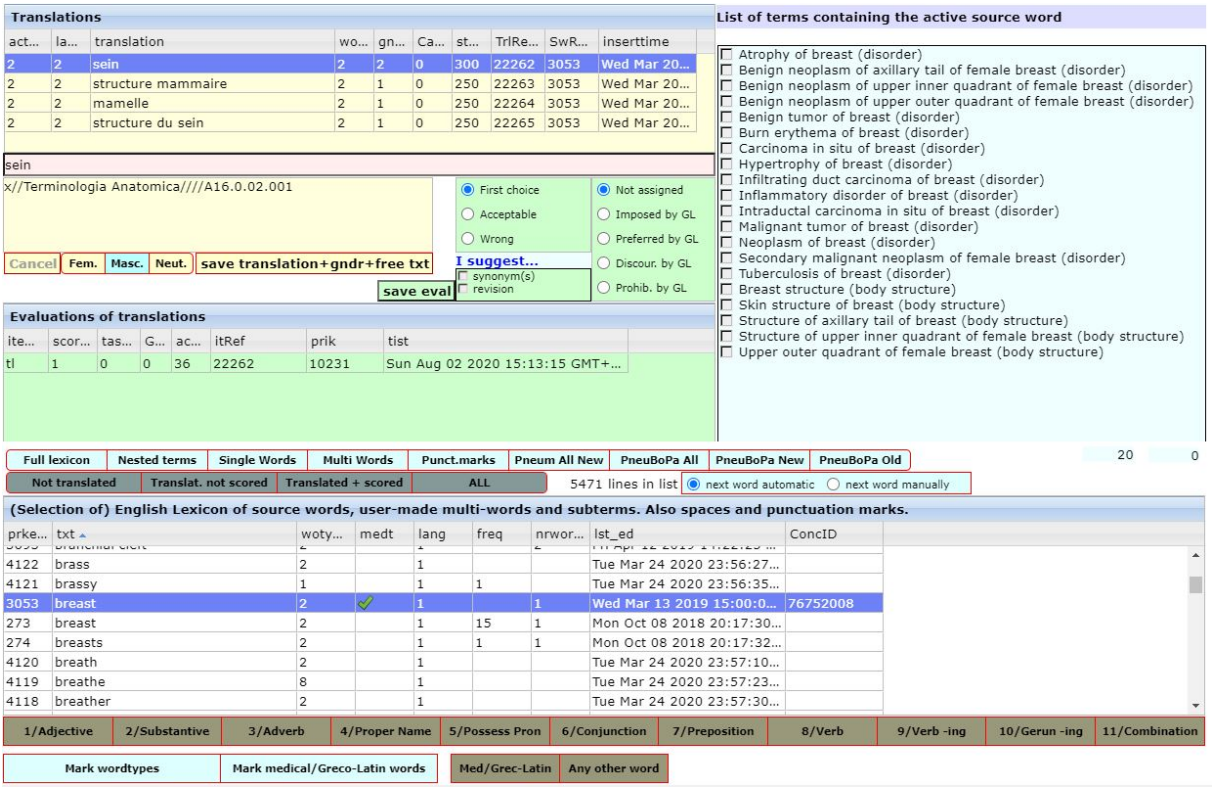

Figure 6. Screenshot of the CoMeTT lexicon translation interface

In Figure 6, the lexicon word "breast" is highlighted in dark blue in the lexicon's alphabetical list. Also highlighted here is the English word type (2=noun) and, in the last column, a SNOMED CT identifier because this word is also in itself a SNOMED CT concept. The top yellow box shows the French translations of this word, including the gender and word type along with other metadata. The top light blue box shows the SNOMED CT FSNs using this term. The various buttons allow the addition of word translations and of metadata on the source word and its translations.

3.3. Roles, functions, methods and decision-making processes

In CoMeTT, the following roles are included: project manager, coordinator, IT and database manager, domain expert translator (DET) and linguist. The translator role is divided into functions. All roles (except Linguist) can also take the role of DET. The functions of DETs are reviewer 1 or 2 (REV1/REV2), or supervisor (SUV). Each DET will be assigned at least one of these functions. A DET (REV and SUV) is assigned one or more batches of source terms to be translated. A DET can have one or more of the three functions (i.e. REV1, REV2 and SUV), but obviously not for the same batch.

The minimum task of REV1 is to rate one translation as the preferred term (PT). This can be either a CoMeTT pre-translation, a translation imported from another project or a new translation created by the REV conducting the rating task. REVs are free to create extra translations for a given concept and rate them as synonyms (SYN). They may also rate translations as "To check" (TC) or "Non acceptable" (NA). When a term is completed, the REV or SUV closes the term. When REV1 closes the term, it becomes accessible to REV2. REV2's task is similar to REV1's, except that the translations and ratings of REV1 will also be presented as pre-translations. SUVs can intervene, either at the request of the reviewer or on their own initiative. SUVs must also check the translations rated as TC by REVs and answer the questions left by REVs in the free text field. When a SUV closes a term, it is closed for everyone (REV1, REV2 and the other SUVs). Only a SUV can re-open a term; it will then be re-opened for all DETs. Like REVs, SUVs can add any translation and evaluate translations already made.

The final decision on which PTs and SYNs to deliver to the NRC is made by the project manager, in consultation with other supervisors, if necessary. Finally, the evaluators (linguists at the NRC) decide on the acceptance of the translations for the Belgian release.

4. Translation of respiratory system concepts with CoMeTT

4.1. The respiratory system translation project (RS project)

Nineteen medical doctors of various disciplines, all members of the CSCT, accepted to cooperate as translating domain experts in the RS project. They were trained in the use of the SNOMED CT ontology and associated terminology by CSCT senior terminologists and in the use of the CoMeTT tool by its developer. The national and international translation guidelines were explained in two live group sessions. The participants were also provided with short CSCT translation practical reference guides. Despite the training and guides delivered, many reviewers expressed a significant need for support regarding both the use of CoMeTT and the terminological work. The translation activities proceeded thus with the close pro- and interactive support of the supervisors.

Even taking into account the impact of the COVID-19 pandemic on each participant, work proceeded at a very heterogeneous speed depending on the batch and on the reviewer. As expected, highly specific concepts like anatomical concepts or very specific disease or procedure concepts (rare diseases, nose-ear-throat surgical interventions, etc.) took the most time to translate, were subject to most questions from the reviewers and required consultation of external references/experts. Again unsurprisingly, reviewers with less medical and/or SNOMED CT experience needed more support.

The acceptance rate of the CoMeTT pre-translations by the REV1 cohort was low. This indicates that machine translations are (still) often incorrect. On the other hand, it does not mean that they are useless. While the word translations were mostly correct, the errors were often grammatical or syntactical. Such errors were easy to correct. However, any correction of a term in the CoMeTT system generates a new entry, and thus it is not counted as acceptance of the pre-translation.

The REV2 cohort changed the PT in 26.4% of the concepts. This included mainly the correction of accidental spelling errors and non-conformance to editorial rules, as well as the application of linguistic templates developed by the SUVs during the revision process in answer to questions from the REVs. Linguistic templates are recommendations regarding the word order, the use of the noun or adjective form and the translation of some words that together provide a frame to ensure that concepts of the same 'family' will be translated in a consistent way across different translators. An example translation template is provided in Figure 7.

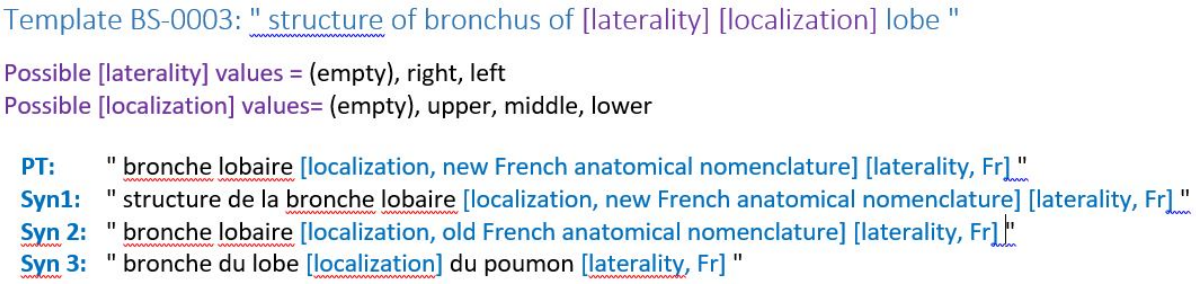

Figure 7. Example of a translation template

SUV corrections on REV2 work mainly consisted in the transversal application of templates, which did not alter the meaning of the REV terms, but enhanced consistency of the wording of the medical terms within the project. SUVs also decided on disagreements between REV1 and REV2, after consulting the literature and/or with external, highly specialised domain experts.

The creation and application of templates went beyond the requirements of translators and users and arose during the project from a spontaneous initiative of the supervisors to further enhance the quality of translation and, ultimately, of the medical language in the EHR.

All classes of DET expressed having spent more time than they had planned on their tasks. This was most apparent in the case of REVs with no previous SNOMED CT translation experience and with SUVs. The concept revision rate increased in all classes of DET as the project progressed, indicating that experience of the tools, rules and process is an important factor influencing speed of translation.

4.2. Efficiency of the approach

Due to the huge amount of SNOMED CT concepts (more than 350,000) and the obvious importance of exact translation of medical concepts, efficiently achieving speed and quality is essential to the translations of SNOMED CT concepts. This efficiency largely depends on the time needed to produce the initial correct translation. Analysis of the AID project (Section 3.1) clearly showed that domain knowledge of REVs is a crucial factor in this context. The experience with the RS project has shown that domain expertise must be interpreted very precisely, also taking into account the different disciplines within a specialty. For example, there is a considerable difference in knowledge between a general practitioner, a pulmonologist and a thoracic radiologist.

The low time efficiency of the RS project had different causes. First of all, the way in which translating domain experts are recruited and the composition of the attributed batches are of the utmost importance in terms of how a group works. In the RS project, the participants were a heterogeneous group of personally recruited medical doctors from various disciplines, and they were assigned random batches of work at the start. This proved insufficient, and better results were achieved once the tasks were re-arranged so that REVs received batches more aligned with their usual medical practice.

Another important point to consider is the motivation of participants, which is decisive for their engagement (Jiménez-Crespo 2017). Gambier (2017: 18) addresses the issue of new types of translators and suggests inquiring into volunteer translators' "motivations, expectations, concepts of translation, their working languages, socio-educational profiles, etc." He concludes that sociological studies such as the so-called "*sociologie des usages"* have not really gone beyond surface studies because the mere quantification of these practices is not sufficient to understand the social position of volunteers, their work, their sociability, the time they devote to the internet, etc. (ibid., emphasis in original). In the present RS project, all participants were connected to the CSCT in one way or another. Loyalty to the organisation and an interest in SNOMED CT translation (as potential end

users) were probably the most important, if not the strongest, motivating reasons. The importance of financial reward for participants is difficult to assess, but it ranged from participants who "do not ask for any compensation" to those who said that the compensation was justified or welcome.

In conclusion, the level of specialisation of the translating domain experts determines how quickly they can correctly interpret certain concepts, which saves them time-consuming research. The most important factor for the fast and correct translation of medical terminology is therefore that the selection of concepts to be translated corresponds to the translator's domain expertise. Motivation to achieve quality results and supervisory support can alleviate an initial lack of experience at the cost of time. Although the CoMeTT tool can maximise the translation speed of the domain experts, it cannot improve their translation quality, unless it provides not just lexicon- but also template-based pre-translations. In the workflow, every step is important. Nevertheless, the first steps—such as the preparation of the lexicon, translation guidelines and templates—are the most important because they heavily influence the following ones. Lastly, our results show that not only the translations, but also the process metadata should be taken into account, as they can provide insights into the structure of the group, the interactions and the output of each translator.

5. Findings and discussion

5.1. Web-based collaboration and communities of practice

During the collaborative translation process, the interaction between the REVs and the SUVs, as well as between the SUVs themselves, was intense and went beyond formal collaboration. A total of 2,256 messages were sent by 15 collaborators, of which 477 were from the REVs and 1,779 were from the SUVs. Of the 1,779, a total of 1,377 came from the SUV project manager. As this is a retrospective study, we need to be careful in analysing the content of the messages for both ethical and scientific reasons. The privacy and integrity of the domain experts was not compromised in the analysis as only high-level, non-identifying features of the message content were described. Concerning the scientific aspect, we can state that all messages were relevant in the context of the project. They concerned either the source term or the reference used (e.g. scientific publication, dictionary) and the concept itself. It is also very revealing that there were a number of messages that added systematic metadata to the translations (e.g. signalling the plural form to a particular editorial template). Finally, there was also interaction by email and phone between REVs and/or SUVs to define new guidelines to adjust the workflow. This interaction not only served knowledge transfer, but also promoted group participation and self-structuring, as described by Yang (2020). In this way, the individual visibility of the participants was increased and an implicit hierarchical

structure was created through the conversational utterances in the free text field (ibid.). As Yang's (2020) analysis of dialogue acts in collaborative environments showed, key actors are essential for motivating a team. In our project, this role was taken up by the SUVs and the coordinator, who guided the team in their collaborative translation efforts, e.g. through face-to-face or email contact.

The fact that all members were highly engaged and willing to share knowledge and learn from each other 'in practice' might lead one to assume that the group gradually evolved into a CoP during the translation process. (See also the description of CoP development as a result of learning processes in practice described in Gherardi *et al.* [1998].) This assumption could also be supported by the fact that all participants were recruited from the CSCT, an open community for the exchange of ideas, consisting of medical doctors, care providers from hospitals or other institutions, linguists, computer linguists and IT service providers who collaborate on the design, translation and implementation of a common terminology model in EHRs in Belgium (https://csct.be). However, despite some common features, a collaborative group differs from a CoP in a number of ways. The main differences between these two forms of organisation have been studied by Storck and Hill (2000) and Lesser and Storck (2001), among others, focusing on three criteria: group orientation, formalisation of relationships and shared practice. Groups work towards a defined goal (e.g. a specific translation task), whereas CoPs define their own purpose. Group relationships are defined by the organisation, whereas CoPs are formed around the practice. Finally, group members share interests and depend on defined processes (e.g. the workflow during a translation process), whereas CoP members share the practice: they are practitioners from different fields and define their own processes that give shape to their practice. Sethi (2017: 7) summarises the essential differences as follows:

> A community of practice is different from a work team in that the shared learning and interest of its members keep it together. It is defined by knowledge rather than by an individual task, and exists because participation has value to its members.

In the collaborative group described, the community-building factor of reciprocal learning was present, but only emerged implicitly over the course of the collaboration. Knowledge sharing and transfer, which are the essential features of a CoP to ensure mutual learning, were only important insofar as they supported the completion of the common translation task. While the group in question was a hierarchically organised team focused on a formally organised work process and predefined work packages with the goal of accomplishing a well-defined translation task within an agreed time frame, "the notion of a community of practice does not primarily refer to a 'group' of people per se. Rather it refers to a social process of negotiating competence in a domain over time" (Farnsworth *et al.* 2016: 143). Unlike a CoP, the group also disbanded after the task had been completed.

5.2. Interaction between collaborative translation groups and CoPs

Given the complexity of SNOMED CT and the multiple translation challenges, the question is whether interaction between the collaborative translation group and a true CoP formed around medical terminology and ontology translation could support relevant translation activities. Related questions include: Who should establish the CoP and what requirements should such a CoP fulfil?

In order to explore these questions, we first need to specify the criteria used to identify the characteristics of a CoP. According to Lave and Wenger (1991), a CoP has three inherent characteristics: a domain, a community and a practice, the practice being primarily a function of reciprocal learning/teaching in solving common problems. As far as the domain is concerned, a CoP for (ontology-based) medical terminology translation should cover the whole range of domains or disciplines potentially involved in medical ontology translation such as: Health Informatics (especially ontology design), Terminology Science, Linguistics, Translation Studies, Software Development, and Data Analytics. Due to the diversity of disciplines involved, the name of the CoP should be as broad as possible (e.g. "Ontology-based Medical Translation"). The different disciplines should be represented by individuals or groups (e.g. institutes, hospitals), each bringing their own practical applications, experience and knowledge. All members should be interested in sharing knowledge with experts from the same and other domains. Medical students and students of Translation Studies and Computational Linguistics could also join the CoP as peripheral members (e.g. for correction of the lexicon or advice on translation and terminology).

The initiators of a CoP for ontology-based medical translation could come from different disciplines and professions. A necessary condition would be that terminology-related and translation-related activities would be within their interests and field of practice. For example, EHR managers would be well suited to build such a CoP. As far as member selection is concerned, good professionals should be approached, especially if they are interested stakeholders, e.g. EHR managers in a hospital. Linguists and translators would be important members as well, as their knowledge is essential to the successful production of high-quality translations. Other potential members could include medical professionals from various fields, the local NRC and software developers who could help improve translation tools and support the establishment and management of translation quality standards with particular attention to terminological consistency.

For obvious reasons, medical specialists-in-training (who are a good example of learning in practice) would be highly useful members as they are engaged in the theory and practice of the medical field and are up to date with the academic literature and the latest developments in their fields.

Moreover, they are still in learning mode, and terminological activities are relevant for them to enhance their terminological knowledge. The collaboration with specialists-in-training could also help to address the problem that few specialists are willing to participate in translation projects. Last but not least, they represent a large recruitment pool, and most of them will be potential end users of SNOMED CT.

Collaborative translation groups that wish to be supported by the CoP could join as a member for the duration of a project, especially if it started from within the CoP. The group could, in exchange for expert support by the CoP, put their knowledge (such as translations or process metadata) at the disposal of the CoP members (care centres, research institutes, SNOMED International), within the limits of official regulations and agreements. Also, process metadata—such as efficiency parameters or group dynamics—function parameters—such as the relationship between quality and sources used—or domain expertise could be made available to interested members of the CoP for their own projects. This 'satellite-like' relationship between the collaborative translation group and the CoP would be temporary and might or might not be launched by the CoP.

The fact that the translating members of the RS project examined in this paper are all members of the CSCT is another important element. Due to its composition, structure, objectives and functioning, the CSCT already forms a CoP, which means that the translating domain experts who participated in the RS project were familiar with the functioning of a CoP. Given the broader scope of the CSCT (i.e. its activities also include education and support to SNOMED CT implementation), a CoP for ontology-based medical translation should be formed as a separate entity within the CSCT.

When deciding how this CoP should be set up, two options could be considered. The first option or 'natural option' would be to let the 'seeds' in the translation group grow naturally into a CoP by consolidating and refocusing the group's mission from pure translation to supporting collaborative translation groups. The second option or 'organised option' would be to establish a CoP artificially, based on invitations, consultations and agreements with people and organisations that could contribute knowledge and experience to support translation work. In the organised option, commitment and willingness to collaborate would still need to be discussed and agreed upon in advance, whereas in the natural option, these foundations are already in place, which means that only acceptance of the mission to support translation groups would be required to activate the CoP. An important initiative of this natural CoP would be to invite external teachers, with the option for them to become permanent members of the CoP if they wanted to learn from the others, engage and share knowledge. This give-and-take balance would have to be assessed by the teachers themselves and would become the basis for their decision whether to join or not.

Important arguments in favour of the natural option are the emergence in the collaborative group of a sense of joint enterprise and mutual engagement, as described by Wenger (2000). These are essential for the resulting shared repertoire, and thus for the growth and sustainability of a CoP.

One argument for opting for the organised form is that CoPs are "only one of the forms of organisation" and that their "emphasis is therefore more on the notion of 'practice' [...] than on 'community'" (Gherardi *et al.* 1998: 278). This option is also supported by the fact that well-structured "legitimate peripheral participation" (ibid.) is very important for the functioning and sustainability of a CoP. This is a particularly relevant aspect for a CoP for medical translation with its strong focus on training (see Section 5.3). In any case, "collaborative thinking is necessary for CoPs to thrive" (Iverson and McPhee 2008, cited in Pyrko *et al.* 2017: 395). As the collaborative group has already developed a certain level of shared thinking and sense of community during the RS project, starting with this group probably offers the most guarantees for success and sustainability. The risk of failing with an organised option is well illustrated in the study of Pyrko *et al.* (2017: 402) with the example of leaders who "tried to 'set up' a CoP by focusing on the tools but neglecting the organic nature of CoP development".

Members of such a CoP could support collaborative translation groups through different initiatives, such as: planning or selecting translation projects based on specific needs of the medical field and recruiting subject experts; organising training in health informatics for translation students or CoP members from other disciplines; monitoring and improving ICT support; analysing project outcomes; promoting collaboration and exchange of data and experiences, e.g. in the form of seminars or conferences.

An example of an established CoP in the context of SNOMED CT is the Translation User Group (https://confluence.ihtsdotools.org/display/TRANSLATIONUSERGROUP), which has the mission to support SNOMED International's strategic objectives by advising on translation activities, in particular by setting guidelines for translation quality assurance and translation process management. The Translation User Group provides a forum for the SNOMED International community to share knowledge and discuss translation issues related to SNOMED CT, including suggestions on priorities for translation activities, best practices for implementation and specification of requirements for translation tools. The goals are to gather input from different perspectives on the current state of the art in Terminology and Translation Studies, to identify best practices in relation to translation practice and to support members in their translation activities.

5.3. The potential of collaborative translation groups and CoPs for translation training

Both collaborative translation groups and CoPs could be useful environments for situated learning. The emphasis in translator education today is on developing the learner's competence as a translator (Kelly and Way 2007; Kiraly 2000), which means that training is no longer limited to abstract and decontextualised classroom activities. Rather, students are also exposed to real and/or highly simulated work environments and tasks outside the classroom (González-Davies and Enríquez-Raído 2016). As argued by Brown *et al.* (1989: 33), learning is situated because it is a "function of the culture and activities" in which it takes place. In a process of "legitimate peripheral participation" (Lave and Wenger 1991), learners in situated learning participate in a CoP, becoming experts in an environment where theory is put into practice. There are many different models of situated learning, all of which have the following features in common: apprenticeship, collaboration, reflection, coaching, multiple practice and articulation (McLellan 1994). Within a collaborative translation group, medical translators-in-training and terminologists could, for example, participate in a project within the framework of an internship and become involved in the workflow by performing specific tasks such as reviewing (e.g. automatically generated pre-translations), translating or revising with other group members. In this way, they could gradually become active members of the group. Different types of knowledge and skills could be shared in this process. Students could pass on knowledge about (specialised) language and translation and terminology acquired in class to medical professionals. This would give them more insight into the theoretical and practical aspects of their terminology work, as well as into translation as a professional activity, which would help them to develop good practices (e.g. consistent use of terms or adherence to translation guidelines). In turn, students would acquire knowledge in an authentic context, which would include aspects such as: working with participants from different backgrounds (domain experts, IT specialists) or translating in a practice-oriented way, where linguistic rules and terminological principles have to be carefully balanced (clinical language use often contradicts terminological rules), to name just a few examples. Learning in such a collaborative working environment would be to a large extent unintentional, as it would arise from task performance rather than from abstract learning in formal teaching. This would encourage translators-in-training to think and act like professionals and, not least, to develop and optimise their social skills.

Medical translation trainees could also benefit from a CoP formed around the translation of (ontology-based) medical translation, at least as peripheral members, by acquiring knowledge in other fields related to terminology in one way or another (e.g. health informatics and ontology), which would broaden their view of (specialised) translation and terminology activities in different fields. On the other hand, integrating students in the

CoP would improve its performance when translating SNOMED CT into specific languages. Lecturers in Translation Studies or Terminology Studies could also be active members of this CoP, where they could share their scientific expertise with members of other, non-linguistic fields. Collaborative translation projects with domain specialists, translators-in-training and health professionals-in-training are already at a planning stage, and it is certainly worthwhile to develop this approach further.

6. Summary

This article discussed the usefulness of a CoP to support the collaborative web-based translation of medical ontologies, in particular that of SNOMED CT. We started with a description of the SNOMED CT ontology and the translation tool CoMeTT, specifically designed to support the translation of this ontology. The basic principles of the tool were discussed and the results of an efficiency test were briefly mentioned. The results of a translation project for medical terms of the respiratory system carried out by a collaborative translation group of medical professionals were then analysed. The composition, recruitment, preparation and functioning of the group were described, as well as the interactions of the members within the translation tool or otherwise. We discussed how a need for support was indirectly expressed through these interactions and how members' behaviour and interaction led to the development of features reminiscent of a CoP: a domain, a community and a practice, with a sense of joint enterprise and mutual commitment as drivers and a shared repertoire as an outcome. We compared a CoP with a collaborative translation group in terms of its definition, characteristics, structure and functioning. We then analysed and discussed the most useful approach for forming a CoP to support collaborative translation groups for medical ontologies. We concluded that the right way forward is to build on the emerging CoP characteristics of the current collaborative translation group by empowering the group and encouraging it to shift its mission from translation to supporting collaborative translation groups. The relationship and interaction between the CoP and collaborative translation groups were then discussed. Finally, the general goals of such a CoP were discussed with special attention to teaching in general and specifically for medical and translation students.

References

- **Amin, Ash and Joanne Roberts** (2008). "Knowing in action: Beyond communities of practice." *Research Policy* 37(2), 353-369. DOI: 10.1016/j.respol.2007.11.003

- **Benson, Tim and Grahame Grieve** (2016). *Principles of Health Interoperability: SNOMED CT, HL7 and FHIR*. 3rd edition. London: Springer.

- **Bhattacharyya, Subhra Bikash** (2016). *Introduction to SNOMED CT*. Singapore: Springer Science and Business.

- **Brown, John Seely, Collins, Allan and Paul Duguid** (1989). "Situated cognition and the culture of learning." *Educational Researcher* 18(1), 32-42. DOI: 10.3102/0013189X018001032

- **Ellis, David, Oldrigde, Rachel and Ana Vasconcelos** (2005). "Community and virtual community." *Annual Review of Information Science and Technology* 38(1), 145-186. DOI: http://dx.doi.org/10.1002/aris.1440380104

- **Farnsworth, Valerie, Kleanthous, Irene and Etienne Wenger-Trayner** (2016). "Communities of practice as a social theory of learning: A conversation with Etienne Wenger." *British Journal of Educational Studies* 64(2), 139-160. DOI: 10.1080/00071005.2015.1133799

- **Gambier, Yves** (2017). "Denial of translation and desire to translate." *Vertimo Studijos* 5(5), 9-29. DOI: https://doi.org/10.15388/VertStud.2012.5.10556

- **Gherardi Silvia, Nicolini, Davide and Francesca Odella** (1998). "Toward a social understanding of how people learn in organizations: The notion of situated curriculum." *Management Learning* 29(3), 273-297. DOI: 10.1177/1350507698293002

- **González-Davies, Maria and Vanessa Enríquez-Raído** (2016). "Situated learning in translator and interpreter training: Bridging research and good practice." *The Interpreter and Translator Trainer* 10(1), 1-11. DOI: 10.1080/1750399X.2016.1154339

- **IHTSDO (International Health Terminology Standards Development Organisation)** (2012a). *Guidelines for Translation of SNOMED CT.* https://www.snomed.org/SNOMED/media/SNOMED/documents/IHTSDO_Translation_Guidelines_v2_02_20121211-(1).pdf (consulted 09.05.2021).

- **IHTSDO (International Health Terminology Standards Development Organisation)** (2012b). *Guidelines for Management of Translation of SNOMED CT.* https://www.snomed.org/SNOMED/media/SNOMED/documents/IHTSDO_Guidelines_Management_Translation_SCT_v2_02_20121211-(1).pdf (consulted 09.05.2021).

- **IHTSDO (International Health Terminology Standards Development Organisation)** (2018). *SNOMED CT Starter Guide*. http://snomed.org/sg (consulted 09.05.2021).

- **IHTSDO (International Health Terminology Standards Development Organisation)** (2021). *SNOMED GLOSSARY*. http://snomed.org/gl (consulted 03.11.2021).

- **Ivanovic, Mirjana and Zoran Budimac** (2014). "An overview of ontologies and data resources in medical domains." *Expert Systems with Applications* 41(11), 5158-5166. DOI: 10.1016/j.eswa.2014.02.045

- **Iverson, Joel O. and Robert D. McPhee** (2008). "Communicating knowing through Communities of Practice: Exploring internal communicative processes and differences among CoPs." *Journal of Applied Communication Research* 36(2), 176–199. DOI: 10.1080/00909880801923738

- **Jiménez-Crespo, Miguel A.** (2017). *Crowdsourcing and Online Collaborative Translations: Expanding The limits of Translation Studies.* Amsterdam/Philadelphia: John Benjamins.

- **Kelly, Dorothy and Catherine Way** (2007). "Editorial: On the launch of ITT." *The Interpreter and Translator Trainer* 1(1), 1-13. DOI: 10.1080/1750399X.2007.10798747

- **Kiraly, Donald** (2000). *A Social Constructivist Approach to Translator Education: Empowerment from Theory to Practice*. Manchester: St. Jerome.

- **Lave, Jean and Etienne Wenger** (1991). *Situated Learning: Legitimate Peripheral Participation*. Cambridge: Cambridge University Press.

- **Lesser, Eric L. and John Storck** (2001). "Communities of practice and organizational performance." *IBM Systems Journal* 40(4), 831-841. DOI: 10.1147/sj.404.0831

- **McLellan, Hilary** (1994). "Situated learning: Continuing the conversation." *Educational Technology* 34(8), 7-8.

- **Pyrko, Igor, Dörfler, Viktor and Colin Eden** (2017). "Thinking together: What makes communities of practice work?" *Human Relations* 70(4), 389-409. DOI: 10.1177/0018726716661040

- **Santos, Claudia and Rute Costa** (2015). "Domain specificity." Hendrik Kockaert and Frieda Steurs (eds) (2015). *Handbook of Terminology.* Volume 1. Amsterdam/Philadelphia: John Benjamins, 153-179.

- **Sethi, Aditi** (2017). "Learning in communities: Understanding communities of practice in the development sector." *Knowledge Management for Development Journal* 13(3), 4-21.

- **Spackman, Kent A. and Guillermo Reynoso** (2004). "Examining SNOMED from the perspective of formal ontological principles: Some preliminary analysis and observations." Udo Hahn (ed.) (2004). *Proceedings of The KR 2004 Workshop on Formal Biomedical Knowledge Representation*. BC Canada: Whistler, 72-80. http://ceur-ws.org/Vol-102/spackman.pdf (consulted 24.06.2021).

- **Storck, John and Patricia A. Hill** (2000). "Knowledge diffusion through 'strategic communities'." Eric L. Lesser, Michael A. Fontaine and Jason A. Slusher (eds) (2000). *Knowledge and Communities.* Boston: Butterworth-Heinemann, 65-83. DOI: https://dx.doi.org/10.1016/B978-0-7506-7293-1.50008-1

- **Wenger, Etienne** (2000). "Communities of practice and social learning systems." *Organization* 7(2), 225-246. DOI: 10.1177/135050840072002

- **Yang, Jun** (2020). "Participatory, self-organising, and learning: The patterns and influence of peer communication in online collaborative translation." *Target* 32(2), 327-357. DOI: 10.1075/target.19156.yan

Biography

Dr Maria-Cornelia Wermuth (PhD, Vrije Universiteit Amsterdam) is an Associate Professor at the Faculty of Arts of KU Leuven and Head of the Research Group Translation and Technology. She teaches courses of German Linguistics, Terminology and IT and Medical translation. Her research areas are medical translation, terminology, translation tools, terminological computer applications and applied cognitive linguistics. She is an active member of the SNOMED International Translation User Group.

E-mail: cornelia.wermuth@kuleuven.be

Dr Michel J. Walravens (MD, KU Leuven) is a rheumatologist. As a consultant at the University Hospital of Leuven for clinical and laboratory diagnostics of systemic autoimmune diseases, he was active in education, research and social patient organisations. Since 1990 he developed his own electronic health record, which led to his interest in SNOMED CT, for which he developed a special translation tool (CoMeTT).

E-mail: mjf.walravens@gmail.com

Dr Marie-Alexandra Lambot (MD, Université Libre de Bruxelles) is a pathologist. She spent several years working in research and teaching before joining the Federal Public Health Service to work on medical data registration. She is currently adjunct to the medical direction of the CHU Saint-Pierre (Brussels), and the coordinator of the Belgian Community Support for Clinical Terminologies (CSCT), which she founded in October 2016.

Email: marie-alexandra.lambot@stpierre-bru.be