Integrating technology in Latvian translation education: untranslated medical terminology management practice using online resources and computer-aided translation tools
Gatis Dilāns, Centre for Applied Linguistics, Ventspils University College

ABSTRACT
This case study describes a terminology management project for translation students that used a weblog, terminological resources available on the internet and Computer-aided Translation (CAT) tools. The project participants (n=15) were students on a professional translation programme at a university college in the Baltic State of Latvia. During the project, students learned how to find medical terminology translations from Latvian into English that were not present in major Latvian termbases or the Tilde Dictionary. In addition, they were able to manage the terminology in a shared online environment using a free weblog publishing tool from Google, Google Spreadsheets and Microsoft Excel, ultimately converting their bilingual glossary into a termbase with SDL MultiTerm Convert to be used with SDL Trados Studio. The findings of the project study indicate that in order to identify an equivalent term in English, students predominantly used partial transliteration coupled with the Google Autocomplete search prediction technique. Overall, the participants found the project to be a positive experience.

KEYWORDS
Translation education, terminology management, computer-aided translation.

1. Introduction
According to Alcina (2009), terminology has become part of university curricula all around the world. Terminology courses do not only appear in translation programmes, they are also actively offered in many sub-areas, such as LSP (Language for Specific Purposes), information science etc. In addition, specialised lexis in the form of terminology has steadily been taught in biology, physics, medicine and engineering classes to name just a few areas.

At the same time, there has not been a continuing discussion about how terminology (both mastery and management) is to be taught in 21st century classrooms. Traditional learning theories propose “a mechanical and repetitive learning, in which students retain information in their memories apparently without meaning” producing inert knowledge (Alcina 2009: 3). In this age of various terminology processing procedures (Kremer et al. 2005) and options in the form of term banks, databases and Computer-aided Translation (CAT) tools (Bowker 2003; Bowker and Marshman 2009; Bowker and Fisher 2010; Chan 2014; Garcia 2014), it would be highly useful to reconsider the ways terminology or rather working with terminology (i.e., terminology
management) is taught for today’s technologically savvy translation and interpreting students (see Bowker and Fisher 2013).

This paper presents an exploratory project study of terminology management (Warburton 2014) in a college-level professional translation education setting in Latvia. During the project, students had to process Latvian medical terms that did not have official translations in English (i.e., those not appearing in official Latvian term bases such as AcadTerm) using weblog, terminology databases, shared Google spreadsheets, and SDL MultiTerm Desktop terminology management software. Through this process, students learned to use and share various search, translation decision-making and organisational techniques that enhanced their awareness of how terminology could be efficiently managed for specific purposes such as specialised terminology translation.

2. Background

2.1. Terminology management training in translation education

Terminology as a specialised lexicon is crucial in translation. Already at the beginning of the 1990s, Sager (1990: 5) pointed out that “terminology collection and processing is a semi-automatic process, constantly responding to innovations borrowed from information technology.” Warburton (2014: 648) defines terminology management as an activity involving

> a wide range of tasks focused on terminology data, i.e. terms and information about terms such as definitions, context sentences, and grammatical information. These tasks include collecting, developing, storing, reviewing, harmonising, enhancing, and distributing terminology data.

She also notes that nowadays terminology is normally managed by using computers and terminology databases.

While discussing the standardisation of legal terminology in the Welsh language, Davies (2006) pointed out a dynamic, multi-disciplinary approach involving legal, linguistic and language technology experts. According to him, the legal expert would ensure the accuracy of legal concepts underpinning any terms, the language expert would scrutinise linguistic coherence and quality, with the language technology expert probing ways of handling terms using various IT tools.

Although over the past two decades terminology processing and management has made remarkable technological advances in the form of term banks and terminology management software, there has not been sufficient joint interest from either applied linguists, terminologists or IT specialists as to how those
advances could be integrated in the process of educating future translators. Moreover, little has been done to address possible collaboration among field experts, terminologists, IT specialists and academic institutions with the aim of dealing with the dynamic nature of terminology development in fields such as biology and medicine. Ha (2007), for example, notes that fresh terminological developments in medicine and bio-chemistry occur on a constant basis affecting human lives which calls for more efficient techniques and systems of terminology processing.

Montero Martínez and Faber Benítez (2009) discuss the role of teaching terminology in translation training, an issue that had received little attention up until that time. They point out that the subject of terminology in translation programmes needs to be taught with an emphasis on real-life contexts of processing and managing this type of specialised lexis as communicatively dynamic phenomena. Consequently, new strategies in managing terminology in a professional way have to be proposed with the aim of producing better translations. Those strategies are driven by the realisation that translators often lack reliable terminology sources and cannot find translations of terms. Therefore, as per Montero Martínez and Faber Benítez (2009: 92), would-be translators need to be equipped with the following strategic skills: identification of specific concepts, evaluation of information sources, recognition of concept-based interlinguistic correspondence and information management. All this would compensate the necessity of acquiring expert knowledge in a particular field, provided there is a clear idea of how to extract recurrent linguistic templates in both languages (Faber 2012). As per teaching methodology, the authors suggest group work that includes experimental, task-based, problem-solution terminological exercises while using various text sources and state-of-the-art term management software. Results of the exercises are then to be reviewed, discussed and evaluated.

In addition, Bowker and Marshman (2009) call for an integrated translator training approach where terminology management tools not only become part of some specialised courses, but further are sufficiently used across the board. According to them, this would enable students to appreciate the usefulness of such tools to a much broader extent. The authors point out that due to an almost unfettered availability of computers, electronic texts and specialised software, the field of terminology is a frontrunner as far as the use of technologies are concerned. While specifically focusing on their CERTT (the Collection of Electronic Resources in Translation Technologies) project, they also note that a prevailing number of translators use technologies (such as translation memories, search engines, term banks and terminology management systems) in their work. They also describe what software providers and university translation programmes lack in this respect. Consequently, the need for a broader university-based integrated language
and translation technology training that is not limited to technology courses is self-evident.

2.2. SDL MultiTerm Desktop and online termbases

According to Garcia (2014: 68), CAT systems are software applications created with the specific purpose of facilitating the speed and consistency of human translators, thus reducing the overall costs of translation projects while maintaining the earnings of the contracted translators and an acceptable level of quality.

The value of termbases as CAT tools has already been emphasised by ten Hacken and Parra (2008). This value today goes far beyond mere term extraction from compiled electronic texts or corpora to identify and verify potential term candidates (Bowker and Fisher 2013). SDL MultiTerm Desktop is a terminology management tool that, besides term extraction, also allows integration with the SDL Trados Studio work environment (Alcina 2008; Walker 2014). To put it briefly, it was designed for “searching, editing, creating, and maintaining terminology databases” (SDL 2014).

For the purposes of the project described in this article, the functionality of SDL MultiTerm Convert application has to be mentioned. This application allows the conversion of Excel glossaries into an SDL MultiTerm termbase that, again, can be used in SDL Trados Studio as an additional translational source. Such glossaries compiled and managed either in Google Spreadsheet or TaaS (Terminology as a Service; see Pinnis et al. 2013) shared cloud environments by multiple users allow exported data to be further utilised for work in other CAT environments (such as SDL MultiTerm Desktop and SDL Trados Studio). The combination of shared terminology data management and CAT tools creates a powerful technological work enhancement as far as fast, accurate and consistent technical translation is concerned. These glossaries or lists can be compiled in view of what is either available or not available in existing online termbases or in other sources and managed as a crowdsourced project (Ambati et al. 2012).

Terminology storage, processing and management tools are thought to be one of the first technologies used by language professionals; the first termbase was created in 1963 (Bowker and Marshman 2009). Its name was originally Eurodicautom, but it was later renamed IATE (Inter-Active Terminology for Europe). The updated IATE termbase was launched in 1999 and completed in 2004, with the aim of providing an online platform for all terminology resources that are used in the EU and it includes 1.4 million term entries. This type of termbase was eventually emulated all around the world including Latvia whose terminology development was somewhat hampered by the years of Soviet
domination that, as far as specialised language was concerned, put more emphasis on Russian (Veisbergs 2001).

Currently, there are three major termbases in the Republic of Latvia: EuroTermBank, AkadTerm, and VVC. The largest and most ambitious in terms of its scope is EuroTermBank (Vasiljevs et al. 2008). According to its website (see below), the EuroTermBank terminology management project aims to harmonise and consolidate terminology work done in new EU member states. This is done by sharing experience of the existing European Union terminology networks in order to ensure multilingual communication in such areas as law, trade etc. The problem is that the new EU members’ multiple terminology sources are not consistent enough, lacking coordination in terms of term development and technical compatibility. Consequently, the EuroTermBank presents an attempt at a centralised online termbase for Estonia, Hungary, Latvia, Lithuania and Poland. However, the termbase is also open to other new EU member countries and interested states outside the EU area. It contains more than 200,000 Latvian terms with the total number of terms in all languages around 1.4 million. Finally, the project is coordinated by a well-known Latvian IT company Tilde, which specialises in language and translation technologies and produced the Tilde Dictionary which is the first multilingual electronic dictionary in Latvia (now available both as a software package and an online tool).

One of the project partners in the EuroTermBank Consortium is the Latvian Academy of Science, which also has its own termbase called AcadTerm. As noted on its website (see below), the termbase is compiled under the auspices of the Latvian Academy of Sciences Terminology Commission which coins and confirms term glossaries that are supplemented by terms from other collections. This prestigious termbase includes close to 900,000 terms in six languages that are compiled from more than 70 printed terminology sources. In a way, this project of strong national significance has successfully attempted to digitise printed collections of various sorts of printed terminology dictionaries. There is no specific information on how many Latvian terms the termbase contains, though.

A more recent development in electronic terminology processing and management in Latvia is presented by the VVC termbase (Valsts valodas centrs ‘the State Language Centre’). This online termbase combines its own term collection with the terms that have been confirmed by the Latvian Academy of Sciences Terminology Commission and includes over 352,000 terms (presumably all in Latvian). This official termbase allows users to verify terms through reading various additional comments on sources and the context of the term’s use. It also includes obsolete terms with historical significance for research purposes.
To sum up, the research question of this study, consequently, is as follows:

Is it beneficial to use and integrate diverse technology tools (such as weblogs, termbases and CAT software) while teaching terminology management in a professional translation programme?

3. Methodology

3.1. Project study, participants and setting

This project study was conducted as a part of two terminology courses spanning the period of the spring and the fall semesters of 2013 at a university college in Rīga, Latvia. The courses were mandatory in order to complete a four-year professional college programme in translation and interpreting. The participants in the project study were then the second and third-year students attending the courses \( n=15 \). There were thirteen females and two males, with an average age of 21. All of them had Latvian as their A language and English as their B language as far as translation was concerned (Spanish was language C). During the two semesters, the project study was conducted in two specially designated computer rooms that were equipped with internet access, SDL MultiTerm Desktop terminology management tool and SDL Trados Studio state-of-the-art translation software.

3.2. The project MedTermAngliski+

The project study MedTermAngliski+ (translated as MedTerminEnglish+) used a weblog that was created to provide instructions and serve as a platform to search for and log untranslated medical terms (Figure 1). Both the instructors and the students had access to the weblog which was easily manageable through the blogspot.com service provided by Google.
As Figure 1 shows, the project weblog main page interface consists of a number of stand-alone pages covering project description (Par projektu), a video on how to participate in the project (VIDEO: Kā piedalīties?), source and work method description (Avoti un darba metode), online text sources and a shared glossary document (Teksts tiešsaistē (1) un glosāriji (2)), termbases and translation memories (Terminu datubāzes (3) un atmiņas), CAT pages (SDL MultiTerm Desktop, SDL Trados Studio), student work presentations (Studentu darba prezentācijas) and contacts (Kontakti). In addition, the page also displays a search box (Meklēt medtermiņus ‘Search medical terms’) and an archive of all posts, which in the case of this particular project, total more than 500. This means that as a result of this project around 500 medical terms were found that were not present in major Latvian termbases (described in section 2.2.) and their translations were subsequently generated by student participants. The project data has been published on the project’s blog site (both as a glossary sheet and SDL MultiTerm Desktop termbase).
At the centre of the main page, there is a blog post column that usually indicates the date and topic of the post and which was used to publish term entries. As shown in Figure 2, the weblog entry published by a student includes the date of post, source term in Latvian and its translation in English published as a topic of the post (periostīts = periostitis) followed by a structured mini-report describing search results in the termbases (Meklēšanas rezultātī; zero designates that nothing was found), translation method (Tulkošanas metode; partial transliteration in this case) and a section to input data in the future in case the translation appears in a termbase (Termina parādišanās datubāzēs). The last section is intended as a follow-up activity in a future phase of the project involving another group of students. At the very bottom, the post includes information about those who published the entry and a comment option followed by networking options.

The procedure for this educational project study generally involved direct online search techniques guided by human, not artificial intelligence (which constituted good practice for the student participants). The students were given instructions that asked them (1) to select medical terms in Latvian that, in their opinion, would not have translations in English, using online texts, (2) to verify that those terms really did not have translations in the termbases and the Tilde Dictionary, (3) to try to find the translation by phonological similarity (transliteration), definition comparison, images etc. (Figures 4 and 5), (4) to publish the term and its translation as a blog post for a local search (Figure 2) and to copy the term and its new translation in the shared Google Spreadsheet to compile a bilingual glossary that can be exported to Excel (the spreadsheet...
had to be consulted as well to avoid duplication - Figure 3). In other words, during this collaborative online terminology management project, students and two instructors searched medical publications in Latvian on the internet (e.g. term glossaries, professional journals, drug descriptions, hospital websites etc.) for random terms used in the field of medicine that did not have a translation into English in the major Latvian termbases or dictionaries. These include the EuroTermBank, AkadTerm, VVC, and the Tilde Dictionary. When the above-mentioned sources did not provide the English translation of the term, the translation was generated using various search and recognition techniques (mainly partial transliteration owing to a recognisable phonological form of the term that was already borrowed - see Figures 4 and 5). The project participants were invited to comment on the proposed translations (see Figure 2) or to note cases where there was a translation available.

![Figure 3. Sample of shared Google Sheets glossary of medical terms and their translations](image)

The end product of the project was in the form of a shared Google Sheets glossary (Figure 3) which the participants both consulted and compiled during the project. For the purposes of the course, it was only superficially edited and converted into a Microsoft Excel spreadsheet. Then SDL MultiTerm Convert, which is a part of SDL MultiTerm Desktop terminology management software, was used, allowing the conversion of the bilingual Excel glossary to SDL MultiTerm XML format. The XML format was then imported into a newly created termbase that could be used with SDL Trados Studio. The glossary conversion process was a team effort, but each participant had to verify the operability of
the new termbase in the SDL Trados Studio environment individually. In a nutshell, the procedure was designed to allow students to work on terminology management in a shared environment and then learn to convert larger data sets into termbases through SDL MultiTerm Convert for further use with SDL Trados Studio.

4. Results

4.1. Basic techniques of finding English equivalents for untranslated Latvian medical terms

Several techniques emerged as the participants searched for Latvian medical terms that were not available in the major Latvian termbases or the Tilde Dictionary, but which were used by the Latvian medical community. To start with, the participants tried to find the term definition in Latvian and search for an analogous definition in English. However, the prevailing and most straightforward techniques were either partial transliteration or partial transliteration coupled with Google Autocomplete to search for predictions leading to a successful identification of the term in the English language. As Figure 3 shows, the identified untranslated Latvian terms had already been formed by medical professionals on the basis of their English equivalents (e.g., hemianopsija ‘hemianopsia’). In some cases, they represented complete equivalence (e.g. hemangioma ‘hemangioma’).

**IHTIOZE = ICHTHYOSIS**

- Translitterācija
  - H ➔ ch
  - I ➔ y

starp t un y ir h
oze ➔ osis

Figure 4. Description of partial transliteration and Google Autocomplete search prediction technique.
Figure 4 shows a description of partial transliteration and the Google Autocomplete search prediction technique used by one of the participants in the project (the illustration is taken from her project presentation). In Figure 4, the participant demonstrates how she was able to do a translation for the term *ihtioze* 'ichthyosis’. First, she tried to type in the search box a hypothetical phonological equivalent, knowing that transliterating *oze* to *osis* might produce some results. After that, she realised that the partial transliteration pattern for the term was more complex than she had expected. As it turned out, Latvian *h* had to be substituted with the English *ch* and Latvian *i* with the English *y* or *hy*. The Google Autocomplete search prediction technique was initially suggested by one of the project instructors and it was successfully adopted by the participants in multiple variations. At the same time, the participants displayed resourcefulness in discovering unique search techniques of their own that involved both linguistic and technological expertise and ingenuity.

![Figure 4: Partial transliteration and Google Autocomplete search prediction](image)

**Figure 5. Description of parallel information search in Wikipedia**

For example, another project participant chose to search for parallel text information in Wikipedia. She wanted to find the translation for the term *dzirdes caurule* ‘eustachian tube’. Again, the English translation of the term was not available in the Latvian termbases at the time of the search. The
participant decided to investigate the anatomy of the ear on a Latvian Wikipedia page and found a figure that described it in detail. Ultimately, she found an analogous figure in the English language parallel text section on the same topic. Therefore, as Figure 5 shows, she was able to locate the term’s translation. All in all, linguistic knowledge, partial phonologic transliteration and creative search approaches aided the participants in providing translations for Latvian medical terms that were found to be in use by the medical community, but lacking respective translations in current Latvian termbases.

4.2. Follow-up interview with the project participants

After the project was completed, a structured interview was conducted with six female project participants who were available and who consented to take part in the interview. They were asked questions about their general impression of the project, its benefits, and specific features of it which were more or less successful. The questions were as follows: (1) What is your overall impression of the medical terminology management project MedTermAngliski+? (2) What was it that you gained most from the project? (3) Do you think that the experience you gained in the project will help you to professionalise your translation work as far as terminology management is concerned? (4) What specifically did you succeed in or like in terms of project tasks? (5) What specifically did you not succeed in or like as far as project goals were concerned?

**Summary of responses to Question 1: What is your overall impression of the medical terminology management project MedTermAngliski+?**

The participants noted that the project was valuable not only because it helped them understand what to do in cases where there were no translations available for medical terms, but also how the information gained could further be used in combination with other tools such as SDL Trados Studio. The overall impression was thought to be positive because of the usefulness of the project in creating a website where one can find translations of medical terms that are not available anywhere else.

In general, the project seemed an interesting but time-consuming exercise. Some participants, however, were discouraged at the beginning and did not see the point of the project, but later realised that they could use the techniques learned beyond the programme, i.e., in translation jobs. The participants repeatedly pointed out that the project seemed unusually large and time-consuming educationally, yet at the same time they seemed motivated when they understood that the work they had done was useful not only for them, but also for all people who have problems translating medical texts. Overall, their impression was good.
Summary of responses to Question 2: What was it that you gained most from the project?

Answers to this question were somewhat varied. The participants noted that the project helped them to improve their search skills and gave them experience in learning how to handle difficult translation problems. Through the project tasks, they learned more about the principles of word formation in English, and how a word changes when it is transliterated into another language. They also learned about new termbases and acknowledged that the usefulness of the website they helped to create was in its easy access to terms that could present difficulties during translation. They pointed out that such a website could save time for translators.

Other participants indeed stressed that they had learned various new skills while working with the termbases as well as new solutions for dealing with untranslated terms that were not available in the same termbases. To sum up, they said that they gained understanding and skills in how to search for, predict and manage random medical term translations in a shared online environment.

Summary of responses to Question 3: Do you think that the experience you gained in the project will help you to professionalise your translation work as far as terminology management is concerned?

All in all, responses were positive. Specifically, they tended to suggest that the project seemed to be helpful. The participants learned to work with a number of termbases at the same time. Other participants pointed out that the experience would be directly relevant when translating medical texts where such terminology was included. In addition, the participants highlighted the fact that many medical terms do not have official translations (or any translations for that matter). Consequently, they learned how to deal with untranslated terminology and considered this experience particularly valuable and useful.

Summary of responses to Question 4: What specifically did you succeed in or like in terms of project tasks?

Most of the tasks appeared to be both interesting and exciting to the participants. These included term search and prediction, term management and online publishing in the specially designed weblog. Some participants mentioned that they liked to search for terms using a hypothetical transcription technique that was complemented by Google search engine suggestions. Other participants noted that they liked to examine the meaning of the untranslated
terms and then verify the selected choice. Specifically, they said that the translations of terms that were hard to find constituted the most exciting part of the project.

**Summary of responses to Question 5: What specifically did you not succeed in or like as far as project goals were concerned?**

For some participants, it was difficult to find appropriate term translations in English when they were not available in the existing termbases or dictionaries. The identification of such term translations was made even more difficult by ambiguous definitions in the target language and the possibility of translation variants. Other participants pointed out that there were many terms for which they were not able to identify appropriate target translations at all. They also noted that it was time-consuming to verify whether the term and its translation was in the termbases. They said that it was more difficult to find translations for terms that required additional clarification and for which a partial transliteration technique alone did not suffice. Overall, these terminological challenges were also perceived as a part of the translation process. Although the students did not provide specific examples in their answers, their selected project presentations, which include a number of examples, can be viewed at the project blog site.

5. **Conclusions**

The findings of this project study investigating the benefits of using specific tasks aligned with various technology tools to teach terminology management in a professional translation programme appear to indicate that the participants mastered various search and term identification techniques. They also learned how to manage the workflow the assigned tasks generated, which started with relevant term identification, finding a translation for the term, multiple verification (termbases, weblog glossary) and finally compiling, editing and converting that glossary into an SDL MultiTerm Desktop termbase that was tested in the SDL Trados Studio translation environment. The post-project evaluation interview showed that the participants were satisfied with their learning experience as they identified specific skills that they acquired and challenges they had to overcome in processing untranslated medical terms in the Latvian language. Consequently, the research question can be answered as follows:

It appears to be very beneficial to use and integrate diverse technology tools (such as weblogs, termbases and CAT software) while teaching terminology management in a professional translation programme. This helps to involve students in real-world tasks during which they can learn to find new terminology translations, collect them in a shared online spreadsheet and,
finally, prepare that data to be converted for further use in a CAT environment. The challenge that was generated by the procedural complexity (i.e. having to take many technical steps to produce the end product) only added to the students’ need to work together as a team.

In closing, the project study described here demonstrates how to synthesise various CAT tool capabilities and online terminology resources with translation education objectives while creating a task-driven, collaborative workflow in a college setting; by the end of the course a product that can be individually updated, edited and used in the future was created. In this case, this product was an SDL MultiTerm Desktop termbase that included 500 English translations of medical terms in Latvian that we believe were not available anywhere else. In addition, that termbase was ready to be used with SDL Trados Studio. An interesting future development as far as professional translation education is concerned may include an educational use of the collaborative terminology management platform TaaS that has been coordinated by the Latvian IT company Tilde alongside numerous other European collaborators.

Bibliography


Websites

- AcadTerm
  http://termini.lza.lv/term.php

- EuroTermBank
  http://www.eurotermbank.com

- How to convert a Microsoft Excel file to SDL MultiTerm [Video file].
  https://www.youtube.com/watch?v=APMcWe2vvJI (consulted 10.05.2013)

- MedTermAngliski+ project blog
  http://medtermangliski.blogspot.com

- SDL Trados Studio 2014 Getting Started. Training course.

- Tilde Dictionary
  http://www.tilde.lv/tulkotajs

- Tilde’s Terminology as a Service platform
  https://term.tilde.com

- Tulkošanas izglītības hronika (the author’s blog)
  http://drgatisdilans.blogspot.com

- VVC termbase
  http://termini.vvc.gov.lv

Biography

Gatis Dilāns is Senior Researcher in the Centre for Applied Linguistics at Ventspils University College, Latvia. His current research is focused on language learning, multilingualism and technology applications in translation education. He can be reached at gatis.dilans@venta.lv or gatis.dilans@gmail.com