



**Smart Computer Aided Translation Environment**

**REPORT ON THE RESULTS OF THE ANALYSIS OF  
TRANSLATION SYSTEMS AND PRACTICES OF  
TRANSLATORS**

*Deliverable 5.1*

*March 2015*

Authors:

Iulianna van der Lek – Ciudin, Frieda Steurs, Ken de Wachter  
KU Leuven, Faculty of Arts, Campus Sint-Andries, Antwerp  
Eva Geurts, Donald Degraen, Mieke Haesen, Jan Van den Bergh  
UHasselt, Expertise Centre for Digital Media

# Contents

Introduction .....	4
General Goals of WP5.1 .....	4
<i>Motivation</i> .....	4
<i>Methods</i> .....	4
Study of Translation Systems.....	6
Goals of WP 5.1.1 .....	6
<i>Terminology and definitions</i> .....	7
Methodology .....	8
<i>Literature review</i> .....	9
<i>Selection of TEnTs</i> .....	9
<i>Feature inventory</i> .....	11
Preliminary Findings and Discussion .....	13
<i>Literature Study</i> .....	13
<i>Introduction to six TEnTs</i> .....	26
<i>Modules and features</i> .....	39
<i>Visualisation of TEnT Features</i> .....	68
Conclusions .....	69
Study of work practices of professional translators .....	70
Web survey .....	70
<i>Method</i> .....	70
<i>Respondents</i> .....	70
<i>Conclusions</i> .....	91
Semi-structured Interviews .....	92
<i>Workflow</i> .....	92
<i>Conclusions</i> .....	94
Contextual inquiries .....	94
<i>Participants</i> .....	95
<i>Observational Conclusions</i> .....	113
<i>Personas</i> .....	116
<i>Conclusions</i> .....	118
Final Conclusions .....	120
Acknowledgements .....	121

APPENDIX A: Summary of previous empirical research on TEnTs .....	122
APPENDIX B: Glossary of terms.....	124
APPENDIX C: Web Survey .....	127
References .....	132

## Introduction

The current report documents the results of D 5.1. Analysis of Translation Systems and Practices of Professional Translators that contains two sub-deliverables:

- D 5.1.1.1. Study of Translation Systems
- D 5.1.1.2 Work Practices of Professional Translators

This is the first deliverable of the work package WP5.Workflows and Personalised Interfaces. This work package is a joint research effort between KU Leuven, Faculty of Arts, Campus Sint-Andries, Antwerpen and UHasselt, Expertise of Digital Media.

### ***General Goals of WP5.1***

The goal of WP5.1 is to obtain techniques to analyse translators' personal workflows while using the translation environment. Consequently, these workflows will be used to organize and visualize the features in a way to optimize ease of use and efficiency for translators.

## Motivation

Current translation systems present many useful and powerful features (e.g. side-by-side editor, terminology recognition and alternative translations) and a vast amount of information (e.g. highlighted terminology in sentences, autosuggestions while typing). Although these systems boast many features, not all of them are based on the needs of individual translators. Additionally, some translators may lack technological knowledge or IT skills to properly use these features.

Configuring translation systems is not straightforward and often is a very complex task. Consequently, while all translators would benefit from having a translation system that is organized according to their work practices, needs and preferences, translators with limited experience or computer skills rarely configure their translation system.

## Methods

A user-centered approach has been used to analyse the end user's needs and requirements as much as possible.

### **Task 5.1 Analysis of Translation Systems and Practices of Professional Translators**

Existing translation systems as used by professionals have been studied to obtain insights in the available features of translation systems and work practices of professional

translators in order to know what needs to be considered for the work flows that will be analysed for further personalisation of features.

#### *Task 5.1.1. Study of existing translation systems*

The study of existing translation systems aims at creating an inventory of the various features that are offered by different translation environment tools (TEnTs) in order to support the user needs and requirements analysis in task 5.1.2. The current report presents the preliminary findings of the analysis of six TEnTs, including a brief literature review on the usage of translation technologies and their impact on the translation process. The study will be further consolidated in the course of the project and may be used not only for SCATE but also as a basis for the teaching and evaluation of existing translation environment tools.

#### *Task 5.1.2 Study of work practices of professional translators*

To obtain an overview of different types of translators and different work styles, we have recruited a diversity of professional translators in cooperation with KU Leuven, Faculty of Arts, Campus Sint-Andries, Antwerp and the members of the IAC. To learn about the end users we used such techniques as surveys, interviews, and contextual inquiries. The results of this research will be used as input to create user interfaces that are tailored to the users, their work and the environment or context they work in.

## Study of Translation Systems

### *Goals of WP 5.1.1*

The main goal of the Study of translation systems is to provide a general overview of the different modules and features that are offered by current TEnTs to support the translation and localisation workflows. This inventory aims to support the analysis of the user needs and requirements that will take place in Task 5.1.2.

Our methodology consists of two parts:

(1) A literature study to find out what empirical research has been done so far in the Translation Studies field regarding the ergonomics and usability of translation environment tools.

(2) A feature inventory of six representative translation environment tools (TEnTs) that are: SDL Trados Studio 2014 Pro, memoQ 2014 R2 Pro, Déjà Vu X3 Pro, OmegaT 3.1.8, Wordbee and MateCat 0.4.2.2. We have explored the modules and functionalities of these TEnTs by performing a set of typical translation-related tasks, such as project creation, statistics analysis, alignment, term extraction, translation, and quality assurance checks.

In this report we will present and discuss some of the preliminary findings of our literature review and feature analysis of six TEnTs.

We begin by clarifying some concepts used throughout this study and by introducing our methodology. Note that the methodology concerns only Task 5.1.1. Next section will present and discuss the preliminary findings of our literature review and the feature analysis of six translation environments. We round up our report with a visualisation of the plethora of features identified so far and a description of the next steps.

## Terminology and definitions

In this section we will define some of the most important terms used throughout this report. For a comprehensive but non-exhaustive list we refer you to the Glossary in Appendix B.

Unfortunately, the terminology used in the translation technology field has not yet been standardized and there is still confusion about some concepts. For example, in both industry and academia, there are several concepts used to refer to the tool suite that aid translators in their work:

- *Computer-Assisted Translation (CAT)*: a general term to describe various machine techniques employed to (fully or partially) automate or assist human translation (Feder, 2001:49). Synonyms are machine-aided human translation (MAHT) and machine-aided translation (MAT). There are still doubts in the field whether Machine Translation should be included in CAT as this is mainly referred to as human-aided translation (HAMT) (Feder, 2001, Quah, 2006). Pym (2011:77) argues against the use of CAT because “computers are involved in almost all translation jobs, and in a lot of interpreting as well.” Pym suggests to replace CAT by clear reference to the technologies actually involved (e.g. translation memories, machine translation, terminology database etc).
- *Translation Environment Tool (TEnT)*: a term coined by Jost Zetsche (Allard, 2012) to refer to a software package or suite of programs that provide functions to aid translators in their tasks. This suite of programs may include translation memory, CAT, localization and terminology management tools.
- *Translation Memory System (TM system, TMS)*: the core component of a translation environment tool. It is defined by Lagoudaki (2008:31) as “an application that links a repository in which previous translations and their corresponding source text are stored in a structured and aligned way, so that any new text to be translated is searched for automatically and matched to the available resources associated with the system, in order for the system to be able to suggest a translation.” It is often used as a synonym for CAT or TEnT. Therefore, Zetsche (2014) argues that this term does not serve us well because it mainly emphasizes the translation memory feature, whereas the terminology management system is or should be as important as the TM. Furthermore, the acronym TMS is often used to refer to terminology and/or translation management systems.
- *Translation workstation or workbench* (Melby et al, 1980; Hutchins, 1998; Fulford and Granell-Zafra, 2004) that integrate a series of tools to process multilingual information, OCR scanning, spelling and grammar checkers, access to local or remote termbanks, translation memory and machine translation etc.

Because this research focuses on the entire work environment, including the work set-up, and all the tools and resources that language professionals use may use, we have opted to use the concepts *translation environment* (TE) and *translators’ environment tools* (TEnTs). For the sake of consistency, the same terms have been adopted by UHasselt in their part of the research, WP 5.1.1.2. Note that the term computer-assisted translation (CAT) has

been kept in the translators' survey in Appendix C as this term was used when we launched the survey.

For other terms and their definitions used throughout this report, we refer you to consult the glossary in Appendix A.

## ***Methodology***

Our methodology consists of the following steps:

1. Study literature and industry reports to find out what empirical research and case studies have been done so far regarding the ergonomics and usability of translation environment tools.
2. Set criteria for the selection of the TEnTs to include in our analysis.
3. Select an appropriate evaluation methodology for NLP systems<sup>1</sup> that would help us complete the inventory of features.
4. Explore six representative TEnTs: SDL Trados Studio 2014 Pro, memoQ 2014 R2 Pro, Déjà Vu X3 Pro, OmegaT 3.1.8, Wordbee and MateCat 0.4.2.2. We have explored the modules and functionalities of these TEnTs by performing a set of typical translation-related tasks.
5. Present the preliminary results and discuss the findings.

---

<sup>1</sup> Natural Language Processing



## Literature review

We performed an extensive literature review that includes both empirical papers and industry reports to find out what kind of research has been done in the field of translation technology and its impact on the translators' processes and workflows:

- A collection of empirical research papers that provide an overview of the TEnTs used by translators: Austermuhl (2001), Bowker (2002), Somers (2003), Quah (2006), Lagoudaki (2006).
- Various industry reports and publications published by LT-Innovate<sup>2</sup>, TAUS<sup>3</sup>, GALA<sup>4</sup> and Multilingual Magazine<sup>5</sup>.
- Jost Zetzsche's Translator Tool Box, A computer primer for translators (version 11.0, 2014) and his biweekly newsletter for translators. Both publications provide valuable up-to-date information and TEnT reviews.

## Selection of TEnTs

For the selection of six representative TEnTs, we took two main criteria into consideration:

- The preliminary results of the SCATE questionnaire for translators/interpreters that we launched at the end of last year to get an update on tool usage and workflows.
- The TEnTs used by the translators who participated in our field observations for Task 5.1.2.

Besides these two criteria, we aimed to have one TEnT representing the following categories:

- Focus on freelance translator vs. focus on LSP
- Commercial vs. open-source
- Cloud-based vs. desktop application

This resulted in the following six tools:

- Déjà Vu X3 Pro
- MateCat 0.4.2.2
- memoQ 2014 R2 Pro

---

<sup>2</sup> <http://www.lt-innovate.eu/home>

<sup>3</sup> <https://www.taus.net/>

<sup>4</sup> <http://www.gala-global.org/>

<sup>5</sup> <http://www.multilingual.com/issueDetail.php>

- OmegaT 3.1.8
- SDL Trados Studio 2014 Pro
- Wordbee

## Feature inventory

To create a feature inventory of the TEnT systems and tools that we have selected for analysis, we have followed the EAGLES Evaluation Framework of NLP Systems (1996).

According to the EAGLES evaluation guidelines (Report, 1996), featurization *is a hierarchical feature structure describing an object: its components, functions, attributes and values. It does not require complicating testing procedures and it mainly consists of measuring boolean (presence/absence of feature) and other formal-valued attributes. The featurization testing method can be either based on the user manuals provided by the software developer or checked out by the evaluator himself. The evaluator should have knowledge of the user requirements and knowledge of available translators' tools.*

We used a pre-defined feature list created by GALA LT Advisor and used Boolean (presence/absence of feature) to indicate which feature was present in each tool. We completed the feature inventory list in two ways:

1. By checking the user manuals and online Helps provided by the TEnT providers.
2. By performing a number of translation-related tasks based on our experience as translators' trainers:
  - a. Basic customisations (change layout, adjust fonts/colours, customize keyboard shortcuts)
  - b. Create a translation project in each TEnT
  - c. Create/populate/edit/export translation memories (TM)
  - d. Align bilingual files
  - e. Create/populate/edit/export term bases (TB)
  - f. Extract terms automatically, validate terms (accept or reject), and save them to term base.
  - g. Quality Assurance checks (QA)
  - h. Export bilingual files for external review
  - i. Export target texts (TT) or bilingual documents for delivery

We need to emphasize that the aim of the task analysis was not to test the performance of each TEnT, but to find out what each tool has to offer in terms of features and customisation options.

Since we used only binary and binary nominal scales to indicate the presence or absence of features, our feature inspection method has delivered objective results (Höge, 2002).

To be able to perform the task analysis we prepared a sample kit downloaded from the eCoLoRe<sup>6</sup> website, a platform that distributes materials to support translation technology trainers. The kit contains the following sample files:

### Source Text

- Document Title(s): “Language Transport”
- Filename(s): ctslan\_50A\_en.doc
- File Format: MS Word
- Source Language: EN
- Number of Words: 4,017 (EN)

### Related Material

- Related Text(s): ctslan\_620\_en.doc, ctslan\_630\_en.doc | ctslan\_620\_nl.doc, ctslan\_630\_nl.doc
- Translation Memory: ctslan\_50A\_en-nl.tmx
- Terminology List: None
- Language Pairs Available: en-nl

### Contextual Information

- Sender: SAP (IT systems vendor)
- Recipient: IT systems user
- Text Function(s): Instructive
- Text Type(s): Technical documentation

Content: User manual for SAP’s Language Transport, which is used to import languages into an SAP system

Since a glossary was not provided in the sample set, we manually compiled in Excel a list of terms from EN into NL that we used to check the term base import functionalities of the TEnTs.

Finally, we visualised the most common encountered features in a conceptual map divided per type of workflow that you can find in Figure nr 38.

Note that the current report on Task 5.1.1 presents and discusses some of the preliminary findings of our exploratory study. Since it is work in progress, each section will be further completed and discussed in the course of the project.

---

<sup>6</sup> <http://ecolore.leeds.ac.uk/>

## *Preliminary Findings and Discussion*

### **Literature Study**

In the literature study, we have included both empirical papers and industry reports and other publications:

- A collection of empirical research papers that provide an overview of the TEnTs used by translators: Austermuhl (2001), Bowker (2002), Somers (2003), Quah (2006), Lagoudaki (2006).
- Various industry reports and publications published by LT-Innovate<sup>7</sup>, TAUS<sup>8</sup>, GALA<sup>9</sup> and Multilingual Magazine<sup>10</sup>.
- Jost Zetzsche's Translator Tool Box, A computer primer for translators (version 11.0, 2014) and his biweekly newsletter for translators. Both publications provide valuable up-to-date information and TEnT reviews.

The literature study aims at answering the following research questions:

1. What is the state-of-the art of the TEnT and how they changed the traditional translation process?
2. What do we already know about the TEnT usage and the impact of translation technology on the way translators work?
3. What evaluation methodologies have been used in the past to assess the usability of the existing TEnT?

The findings resulted from the literature review will also benefit WP5 1.2 and WP3.1.1 as both analyse the context in which translators work.

### *State-of-the Art of the TEnT systems*

In order to understand how a translation tool can support the translation process, we first need to understand what translators do, how they do it and why they do it. This will be thoroughly investigated by EDM-UHasselt in Subtask 5.1.2 during the semi-structured interviews with various organizations and Contextual Inquiries with translators at their workplace. However, for this current exploratory study of translation environments, it is useful to take into account the standard definition of the translation process in order to understand how this process has been designed within the TEnT.

---

<sup>7</sup> <http://www.lt-innovate.eu/home>

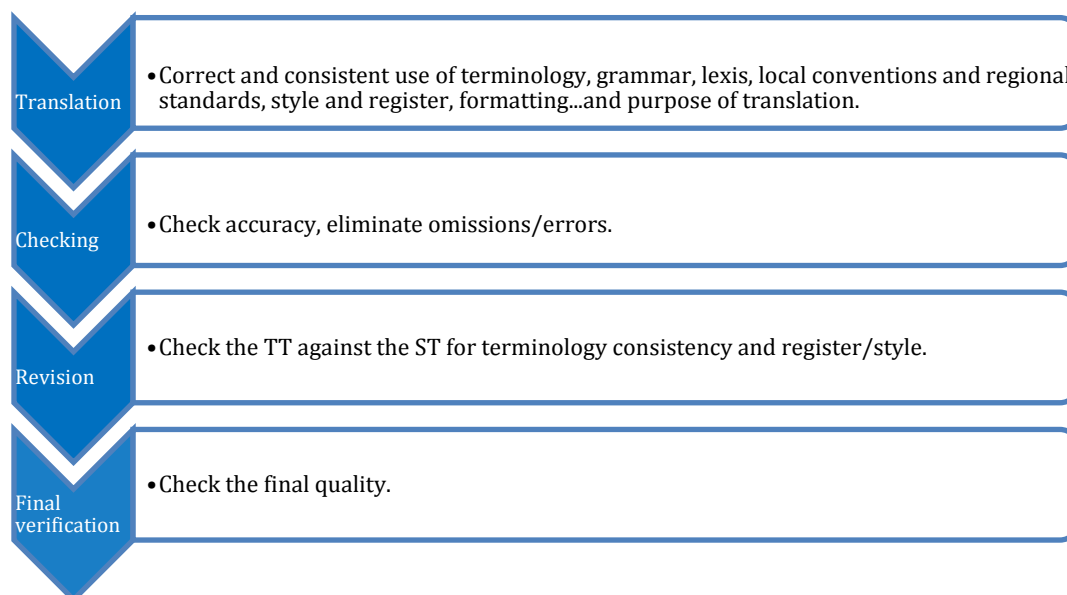
<sup>8</sup> <https://www.taus.net/>

<sup>9</sup> <http://www.gala-global.org/>

<sup>10</sup> <http://www.multilingual.com/issueDetail.php>

## The translation process and technologies

According to the European Committee for Standardization, 2006:12, there are four compulsory stages in the translation process:



*Figure 1 Translation process adapted after the European Committee for Standardization, EN 15038: 2006*

EN 15038 was published by the European Committee for Standardization to ensure the consistent quality of the services provided by the different Translation Service Providers (TSP)<sup>11</sup>.

In this traditional translation process, the workflow consists of four main compulsory stages:

1. The translator receives and translates the source text into the target language
2. The translator revises their own work for accuracy and errors<sup>12</sup>.
3. Once the editing task is complete, another translator checks the entire translation against the source text for terminology consistency and style.<sup>13</sup>
4. Final verification is usually performed by the translation service provider.

<sup>11</sup> Synonym: Language Service Provider (LSP)

<sup>12</sup> Pym (2011) notes that the term 'checking' has not gained standardized status in neither the industry or research and it recommends researchers to use the term "self-revision."

<sup>13</sup> Pym (2011) remarks that the EN-15038 usage of the term "revision" represent neither the industry practices or the research, and hence he recommends using it for both "self-revision" (i.e. "checking") and "other-revision" (i.e. bilingual revision by another person).

Besides the stages listed above, there are two optional stages, namely Review and Proofreading. According to the standard, the Review<sup>14</sup> process is performed by a person other than the translator and it may be monolingual. The Proofreading<sup>15</sup> process is also performed by a person other than the translator and it involves final checks before the target text is published.

With the advent of globalisation, internationalisation and localisation, the traditional translation process<sup>16</sup> has been altered by the use of machine translation and emerging collaborative translation<sup>17</sup> processes where multiple team members work on a single document in real time while the project manager is keeping track of the whole process (CSA report 2007, Garcia 2012).

The translation process and the different workflows are supported by the so-called Translation environment tools (TEnTs)<sup>18</sup> that have started to be commercialised in the early 1990's (SDL Trados, Deja Vu etc). A TEnT has been defined as a software package or suite of programs that provide functions to aid translators in their tasks. This software package may include a translation memory, computer-assisted translation, localization and terminology management tools. Without these tools, the translation vendors would not be able to manage the multilingual digital information.

A comprehensive overview on the state-of-the art of translation technology can be found in Bowker (2002), Hutchins (2004, 2010), Reinke (2013) and in some industry reports and publications published by European Commission (2009), LT-Innovate (2012) and TAUS (2013).

Reinke (2013) remarks that *all in all the core functionalities of commercial TM systems have remained very much the same since the first [...] applications became available at the beginning of the 1990s*. He further comments that the translation workflow has changed dramatically during the last two decades. This is reflected not only in the way the translation processes are organized but also in the way the people involved in the process

---

<sup>14</sup> Pym (2011) remarks that the term is ambiguous because it is also used to refer to any process of revision, book reviews or general proofreading.

<sup>15</sup> According to Biel (2011), this term is not very clear and it should have been further elaborated by the standard. Pym (2011) notes that the term “proofreading” is associated more with editors than with the work of translators.

<sup>16</sup> In localization, the acronym TEP (translation, edit, proof) is being used to refer to the traditional translation process.

<sup>17</sup> According to Pym (2011), collaborative translation may be used as a synonym for crowdsourcing and volunteer translation.

<sup>18</sup> For a comprehensive overview of the electronic tools that have been designed to support the different translation and localization workflows, we refer you to Hutchins' Compendium of Translation Software, online available at <http://www.hutchinsweb.me.uk/Compendium.htm>.

interact and collaborate. Reinke visualises the translation workflow within a TEnT in the following way:

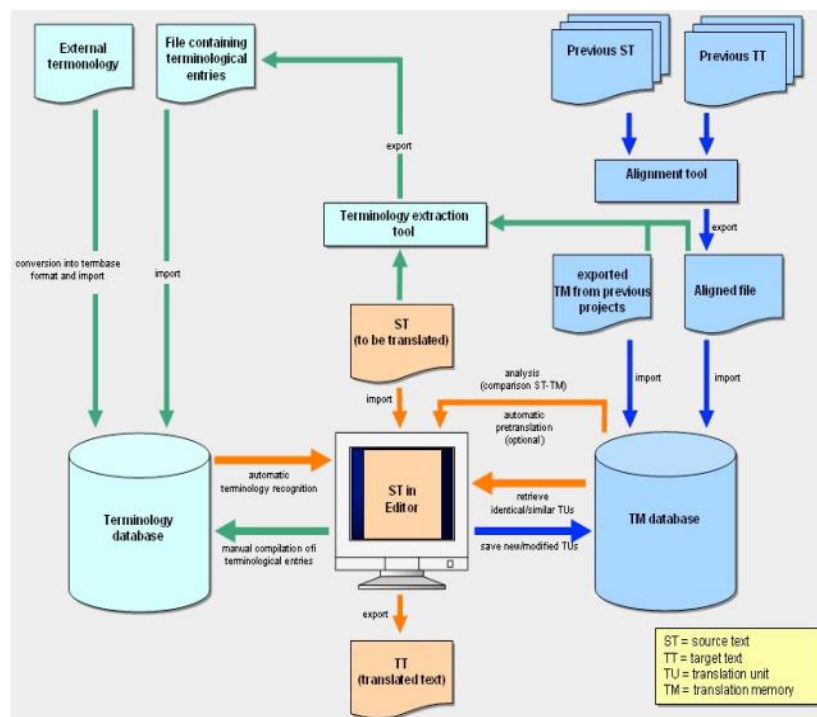


Figure 2 State of the Art in Translation Memory Technology, Uwe Reinke (2013)

In the same line, Mackenzie (2004) and Fernandez-Para (2010) divide the typical translation workflows within a translation environment into three main stages:

- Pre-translation (before)<sup>19</sup>
  - Convert files to a suitable format
  - Align previous translation and import into the TM
  - Extract terms (automatically or manually) to prepare terminology collections relevant for a specific project
- Translation (during)
  - Perform lookup operations in the client-specific TM and TB
  - Confirm/update translations

<sup>19</sup> Not to be confused with the Pre-translation batch feature within a TEnT.

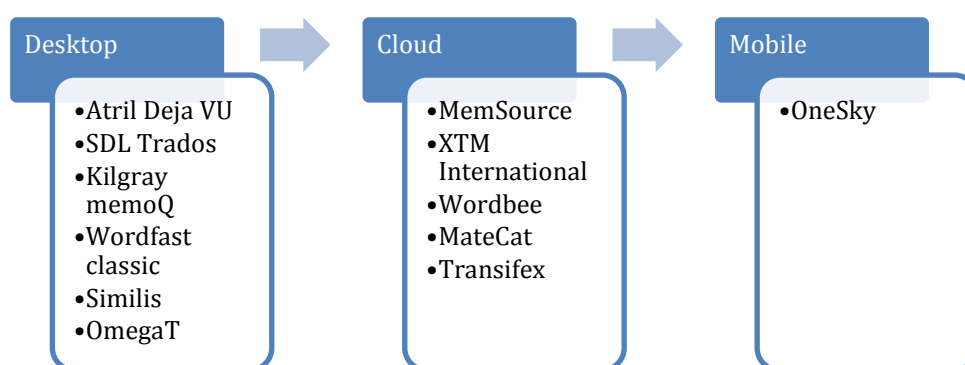


- Post-translation (after)
  - Perform Quality Assurance checks
  - Update resources
  - Export target or bilingual files

Note that both Reinke and Fernandez-Para' workflow representations exclude the interaction with Machine Translation engines and external linguistic resources, such as translation memory and terminology databases.

## TEnT Categories and Usage

There are TEnTs that the user can install on their computer or their local server (client/server-based) and others can be accessed via the browser with a login and a password. Furthermore, there is a new generation of tools with a minimalistic design that can be used on smartphone and tablets for the translation of software strings, e.g. OneSky. However, the latter tools are far from mature yet.



*Figure 3 Examples of TEnTs*

With regards to the TEnT usage, there have been several worldwide surveys conducted by both Academia and the Industry that we summarize below:

Lagoudaki (2006) conducted a large-scale questionnaire survey that reached about 699 language professionals from 54 countries. According to the results of this survey, most used TEnTs by freelance translators were Trados, Wordfast, Déjà Vu, SDLX and Star Transit. In terms of user satisfaction, Deja Vu scores the highest rate, followed by Wordfast and SDL Trados 2006. However, Wordfast gained higher satisfaction rate for usability, learnability and value for money. The same study revealed some of the advantages of using translation memory technology, namely time savings, terminology consistency and improved translation quality.

Between 2010 and 2013, there were three non-academic surveys, Lossner (2010, 2013) and Proz.com (2013) that revealed that SDL Trados, memoQ, Deja Vu, and Wordfast were still the most employed TEnTs. Besides these commercial tools, the open-source tool, OmegaT has gained in popularity.

The cloud-based solutions are praised for the following advantages (TAUS report, 2013):

- The flexible SaaS (software-as-a-service) licensing models

- They enable collaborative translation, by allowing several people to work on the same project in real time.
- Customisable workflow management
- Integration with Content Management Systems (CMS)

Despite their advantages, the cloud-based tools are not yet used on a large scale. A recent unpublished survey conducted by Proz.com<sup>20</sup>, revealed that only 10% of the 1,302 freelance translators who answered the survey, participated in online projects using XTM Cloud. Most used solutions were SDL Trados server (24.06%) and memoQ server (15.09%). Moreover, 44,39% of the surveyed language professionals, have not shared a translation memory. Most concerns for not using an online solution were the low quality of fuzzy matches from other translators and not being able to keep a TM for themselves.

Within SCATE, we launched a general questionnaire at the end of 2014 to get an update on the tools and resources translators are using. The results of this survey are described in the second part of this deliverable, namely Study of work practices of professional translators.

---

<sup>20</sup> <https://www.abroadlink.com/onlineTMsurvey.pdf>

## From classical TM to ‘enhanced’ TM

Due to the recent technology developments, the TEnTs have to possibility to integrate via API<sup>21</sup> technology with various MT engines (both commercial and customised) and external terminology databases (e.g. IATE, Euro TermBank), and parallel corpora (e.g. TAUS Search, MyMemory). These integrations aim at enhancing translators’ productivity, efficiency and consistency.

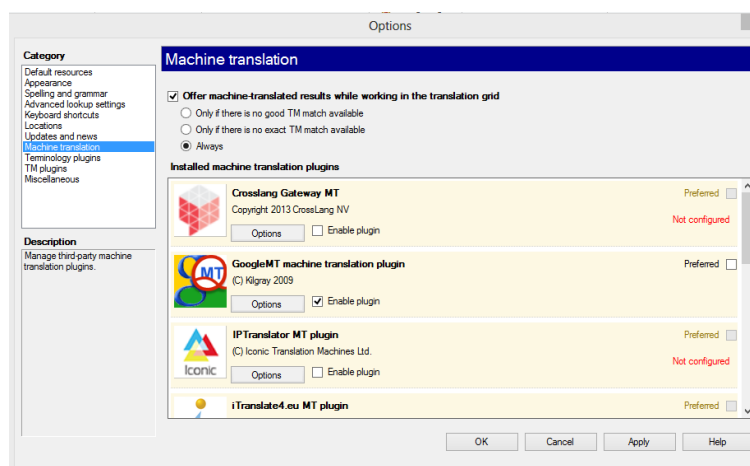


Figure 4 MT options in memoQ 2014 R2

Various experiments both in the Research field and the industry at Audodesk (Plitt & Masselot, 2010) have shown that the TM/MT integration can indeed increase translators’ productivity provided that translators are experienced and they have a positive attitude towards MT and Post-editing (Lange/Bennett, 2000; Guerberof 2008, 2012).

In recent experiments with the students of the Translation Technology Program at the Faculty of Arts, Campus Sint-Andries, Antwerp, we introduced a TM with MT output into the students’ translation workflow with SDL Trados Studio 2014. We found that most students found the TM/MT integration useful for terminology and collocations of technical terms.

Vieira and Specia (2011) investigated a number of TEnTs (SDL Trados 2009, Déjà Vu X2, Wordfast and OmegaT) to find out what kind of features they offer translators to post-edit the matches suggested by the MT engines. Their study showed that the TEnTs under investigation still lacked a number of features desirable for a proper interface for post-editing MT output.

O’Brien and Morkens (2013) also remark that current TEnTs have not been designed to support the MT post-editing task and launched a survey of translators to find out what features would be desirable to optimise the interface of the translation editor for processing

<sup>21</sup> Application programming interface that specifies a set of functions or routines that accomplish a specific task or are allowed to interact with a specific software component. (Source: Wikipedia)

the MT suggestions. The results revealed “machine learning” as one of the most important desirable feature. Such feature would prevent the translators’ frustration from fixing the same MT error over and over again.

As far as the integration of the TEnT with the large online TM and terminology databases, there is little empirical research of their usefulness. The preliminary results of our field research regarding translators’ terminological strategies of acquiring domain terminology (WP3.1)<sup>22</sup> show that some translators found the translation suggestions coming from Euro TermBank useless for the legal type of texts and they preferred to switch off the plugin. Moreover, some of the translators working with Wordbee, found the window displaying the results from the IATE database too small, and hence they preferred to check the database online. Finally, according to a recent webinar on Translation Technology led by Jost Zetzsche (Tool Box Journal, 231st edition), translators would benefit more from the external resources if there would be more advanced ways of pre-filtering the data.

---

<sup>22</sup> SCATE WP 3.1: Translators’ strategies of acquiring domain terminology.

### *Ergonomics and usability*

Technologies combined with the business pressure have led to the so-called “collage translation” (Mossop, 2006) which seems to have an impact on translators’ cognitive abilities.

For example, O’Brien (2006, 2008) investigated translators’ cognitive effort during post-editing TM and MT matches. Their study reveal that the editing of the fuzzy TM matches may increase translators’ cognitive effort and that fixing an MT match has the same impact on the cognitive effort as a 80–90% fuzzy match.

Garcia (2006, 2009), Mossop (2006), Christensen and Schjoldager (2010), and Christensen (2011) all acknowledge that more research is needed on the ergonomics and usability of translation tools.

Lagoudaki (2006) and O’Brien (2010) emphasised that the UI of translation tools has been technology driven, rather than user-driven and therefore, “more usability studies are needed to find out which tools are more productive or user-friendly on the basis of translation interaction with the applications and end-user interaction with the translated text” (Garcia 2006).

Regarding the human-computer interaction topic, the studies of Christensen & Schjoldager (2010), Gouadec (2007), Wallis (2006) distinguish several ways in which translators use the TM technology:

1. The interactive way – when the translators translate one segment at the time
2. The pre-translation mode – when the translator selects the pre-translation operation to automatically pre-fill the target segments with the matches (exact, fuzzy or fragment assembly) found in either the TM, term base (TB) or corpora. If there are no matches provided by the TM and TB, the translator also has the possibility to pre-translate the empty source segments with the help of Machine Translation engines that may integrate in the TM system via API.
3. The web-interactive translation mode – when translators interact with an online TM and they can collaborate with other colleagues working on the same project.

Wallis (2006) found that translators seemed to prefer working in the interactive mode rather than in the pre-translation mode. Garcia (2006) suggested that it would be interesting to compare the productivity and the user-friendliness of the pre-translation mode against the web-interactive mode.

Between 2008 and 2014, we see an increasing interest in ethnographic studies conducted at translators’ workplace: Désilets et al. (2008), Karamanis (2010), Le Blanc (2013), Massey & Ehrensberger-Dow (2014).

During the MT Summit in 2009, Alain Désilets, Research Officer at NRC<sup>23</sup>, Canada, presented a tutorial on the usefulness of the Contextual Inquiry, an HCI technique, during field research involving translators (Désilets et al., 2009). The tutorial was based on his experience assimilated during the research that he conducted at LOPLT<sup>24</sup> that aims at better understanding current work practices of translators, and how new technologies could best support them (Désilets et al. 2008).

Le Blanc (2013) describes the results of an ethnographic study that investigated the translators' perception on the TM technology. Le Blanc spent about 300 hours at three different translation firms in Canada during which he conducted interviews with everybody involved in the translation process and observed translators at their workstations (shadowing technique). As he himself states, the ethnographic approach turned out to be a useful research method as it allowed him to focus not only on the translation processes and workflows, but also on the translators and their work environment. We summarize some of the findings below as they may be relevant current research:

TM advantages	TM disadvantages
Improve consistency, provided that they are continuously updated.	Errors in the TM propagate errors.
Eliminate dull and repetitive work (e.g. updates of user manuals).	The overuse of TM technology hampers creativity and may erode translator's skills over time.
Useful as a searchable database and can help translators acquire new domain knowledge.	May decrease the need of consulting other resources, e.g. dictionaries, term bases.
Make team work much easier when shared in real-time.	Translators may over-rely on the TM matches without too much questioning, which may affect the final quality of their translation.
Allow translators focus on semantic units.	Segmentation affects the whole mental process and thus reduces translation to a mere replacement activity.

*Table 1 TM advantages/disadvantages based on Le Blanc (2013)*

<sup>23</sup> The Institute for Information Technology of the National Research Council of Canada.

<sup>24</sup> Laboratoire d'observation des pratiques langagières technologisées.

Another interesting research project focusing on capturing the translation processes of language professionals at their workplace is the ErgoTrans project initiated by Maureen Ehrensberger-Dow and Garey Massey from the Zurich University of Applied Sciences. This project started in 2013 and it is expected to end in June 2015. Major findings related to the cognitive and physical ergonomics of translators reveal inefficient resource and desktop management, insufficient knowledge of tool features, and ineffective interaction with user interfaces (Ehrensberger-Dow & Massey, 2014). The researchers conclude that translators' efficiency can be optimised by raising awareness about their problems, and by educating them on tools' and proper selection and manipulation of resources.

Based on a classification proposed by Christensen and Schjoldager (2010), we summarized the research on the translation memory technology that took place between 2006 and 2012 in a table that you can find in Appendix A.

Previous research is relevant for WP 5.1.2 and WP 3.1 as they also observe translators in context, namely the aspect of human-computer interaction and translators' strategies of acquiring domain terminology.



### *Evaluation Methodologies of the TEnT Systems*

Although TM systems have been on the market for more than 20 years, little research has been done on their usability. Gow (2003), Simrad (2012), Baldwin and Tanaka (2000), Whyman and Sommers (1999) may be useful as a reference.

O'Brien et al, (2010) study “Keeping an eye on the UI design of Translation Memory: How do translators use the “Concordance” feature?” is relevant for the current study as it reveals some usability issues. Her study is motivated by the fact that “the potential productivity and quality gain from sub-segment matches in TM is not fully realised and may be enhanced by improving the UI design derived from focused research on user experience”. O'Brien's findings reveal that although concordance and AutoSuggest features are useful, translators prefer to have the possibility to close the concordance results window and the AutoSuggest feature off. Similar to Lagoudaki (2006), O'Brien concludes that UI design has been driven by the needs of the translation client and not by the needs of the translator.

However, this attitude seem to have changed with the development of the new generation of enhance web-based TEnTs that integrate both MT and TM technologies. For example, in the development of the MateCat and CASMACAT workbenches, users have been involved in various usability studies: Elming et al. (2014), Alabau et al. (2014), and MateCAT deliverable reports.

The thesis of Tuominen (2012) is also relevant for the current study of translation systems. Tuominen investigated the usability of Wordfast Professional by conducting a usability test in which a group of students carried out a test translation using Wordfast Pro and then completed a user-survey satisfaction. The students' actions on the screen were recorded. Although Tuominen's study has limitations, it revealed some usability issues related to the design layout and functionalities of Wordfast Professional which could have been avoided if the users had been involved in the design process in the development phase. Regarding his research methods, Tuominen concludes that the screen recordings, though useful, turned out to be time-consuming to analyse.

## Introduction to six TEnTs

In this section we will describe the types of interfaces and features that we have identified in the six TEnTs during our exploratory study so far. The aim of this analysis is not to provide a positive or negative evaluation on the selected TEnT systems, but to give an overview of what they have to offer.

Before we describe the various modules and their features, we will briefly introduce each TEnT, present their general user interface and briefly describe the steps we undertook to create a project in each tool.

### *SDL Trados Studio 2014 Pro*

SDL Trados Studio 2014 Pro is a client/server-based TEnT, commercially available since 2005, and it is used by more than 200,000 language professionals worldwide. We have included it in our analysis because it is still a market leader and it offers a plethora of functionalities.

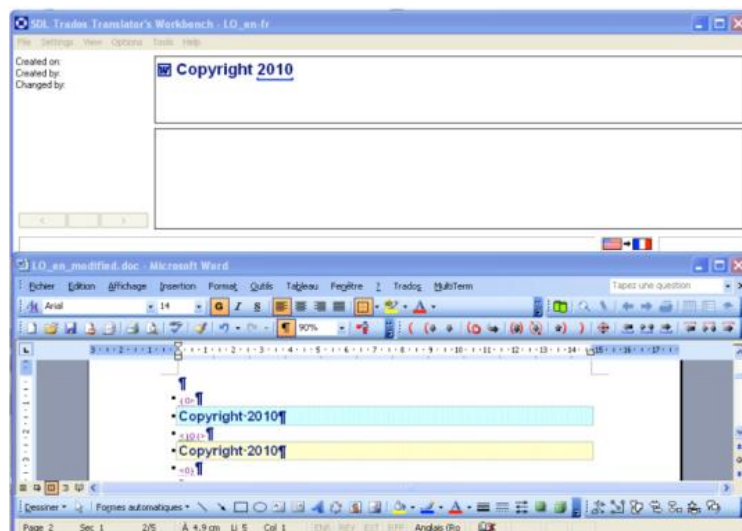
SDL Trados was originally developed in 1984 by the German company TRADOS (TRANslation & DOcumentation Software) as a Language Service Provider (LSP). The company was founded by Jochen Hummel and Iko Knyphausen in Stuttgart. In the late 1980s, the company began developing translation software (early versions of Translator's Workbench, MultiTerm, TAlign, etc.). In 1997, Microsoft decided to use TRADOS for its internal localization memory store and acquired a share of 20% in TRADOS. In 2005, SDL acquired Trados and expanded the functionalities and features for their customers. Earlier versions of SDL Trados Studio 2014 are: SDL Trados Studio 2009 and SDL Trados Studio 2011<sup>25</sup>.

For our investigations, we used an academic licence of SDL Trados Studio 2014.

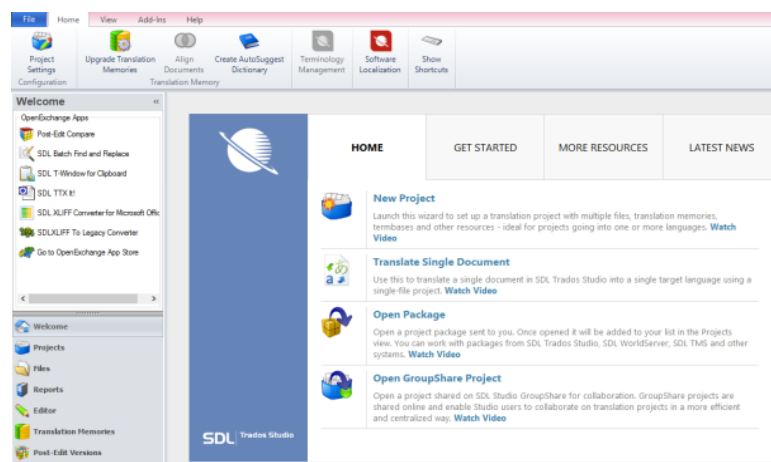
The User Interface has undergone lots of changes since the SDL TRADOS Workbench version in 2007. Figures 5 and 6 show the main UI of the Trados Worbench 2007 and the SDL Trados Studio 2014.

---

<sup>25</sup> <http://www.translationzone.com/about/index-tab2.html>



*Figure 5 SDL Trados Workbench*



*Figure 6 SDL Trados Studio 2014*

From the Welcome page, the users can quickly select the type of workflow that suits them depending on the type of job they have:

- Create a new project
- Translate a single file
- Open a project package sent by somebody else
- Open an online project created in GroupShare.

Through the navigation buttons, the users can access other interfaces, such as the Projects overview, Files, Reports, Editor or Translation Memories.

As far as the UI is concerned, each module is composed of a Ribbon menu at the top, Navigation pane and buttons on the left side, and a working pane where the users can view

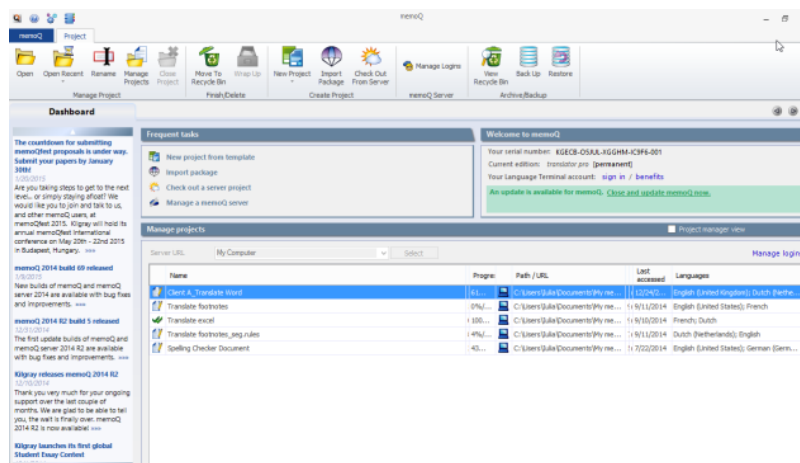
the list of projects and files, or reports. Each window can be docked/undocked or auto-hidden. The users can also switch between windows and even move windows by dragging them to the desired location.

To explore the features in SDL Trados Studio 2014 we chose the *Create a new project* workflow based on the default template. Below we list all the steps and the options the user is presented with during project creation.

- Enter project details
- Add source and target languages
- Upload files for translation
- Create empty TM. Here the user has the option to specify the field settings for the TM and the language resources (e.g. variable list, abbreviations, segmentation rules)
- Once the TM has been created, the user may populate the TM by importing a .tmx file. During the import process, the user has the option to adjust the .tmx import options, e.g. edit the filter, exclude language variants, and overwrite existing translation units (TUs).
- Select and configure Automated Translation engines. When adding the translation memory and automated translation, the user may specify whether they should be used for lookup, concordance searches or updates.
- Add term base provided that one has been created in MultiTerm beforehand.
- Select SDL PerfectMatch. The user has the option to add a previously translated version of the same document and indicate whether PerfectMatch should be applied to the hits coming from that previous version.
- Select Task Sequence, e.g. Prepare without project TM or Prepare
- Review the project settings and click Finish.

### *memoQ 2014R2 Pro*

MemoQ 2014R2 is a client/server TEnT that was launched in 2004 by Kilgray Technologies, based in Hungary. The desktop versions are memoQ Translator Free, memoQ Translator Pro and memoQ Project Manager. For this study, we used a license of MemQ 2014R2.



*Figure 7 memoQ 2014 R2– Dashboard*

The dashboard presents the user with the following workflow options:

1. Create new project from template (Content connector, One Tm and one TB per language pair or One TM per client template)
2. Import package workflow
3. Check out project distributed via the server

To explore the features in memoQ 2014 R2, we chose workflow nr. 1: *Create new project from template*.

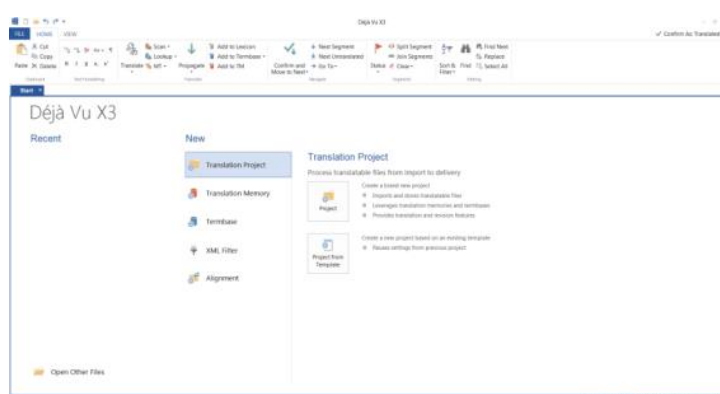
In this workflow, the user is prompted to perform two actions, namely, upload the source files and select the project template. Once the relevant project information (e.g. project name, languages, client name, subject, domain, and deadline) has been added, the user can start translating. Based on the type of template selected, memoQ automatically creates 2 empty TMs (one working and one Master) and one empty TB. When the project has been created, the user may choose to populate the resources, e.g. by importing a .tmx file into the TM, upload monolingual and bilingual files to LiveDocs or importing an Excel glossary in the term base.

### *Déjà Vu X3 Pro*

Déjà Vu was first released in 1993 and it was designed by a translator for translators.<sup>26</sup> Lagoudaki (2006) revealed that Déjà Vu had a higher rate of user satisfaction compared to SDL Trados and Wordfast.

The very first version of Déjà Vu was launched in 1993 to fulfil Atril's own need for a professional translation tool. It was the first translation tool for Windows 3.1 and provided the first interface for Microsoft Word for Windows 2.0 defining Déjà Vu as a very advanced translation tool of its time. Since working with the Word interface imposed too many limitations, Atril began to develop its own integrated environment in 1996. In 1997, Déjà Vu was the first Translation Memory (TM) tool available for 32-bit Windows. It provided its own single translation grid editor and shipped with an extensive number of filters for desktop publishing (DTP) packages, including FrameMaker, Interleaf and QuarkXPress.

There are four editions: DVX Editor (free of charge), DVX Standard, DVX Professional and DVX Workgroup. The latest major release, Déjà Vu X3, was launched in February 2014. For our study, we downloaded a 30-day trial version of Déjà Vu X3 Pro.



*Figure 8 Déjà Vu X3*

The main UI is intuitive and uncluttered, which allows the users quickly access the desired interfaces to perform the following tasks:

- Create a new project or based on a previous template.
- Create a translation memory.
- Create a term base based on the pre-defined fields of Déjà Vu or other templates.
- Create an XML filter to define translatable content.
- Perform and edit an alignment.

<sup>26</sup> <http://www.atril.com/content/discover-d%C3%A9j%C3%A0-vu-x3>

In Déjà Vu X3 the steps to create a project are as follows:

- Enter project name and select project path
- Select the source and target languages
- Create a TM with extension .dvmdb
- Create term base .dvtdb
- Specify MT provider: select and configure engine
- Specify client and subject (optional). The user is presented with a pre-defined list of subjects.
- Upload the source files for translation and adjust the filter options if needed.

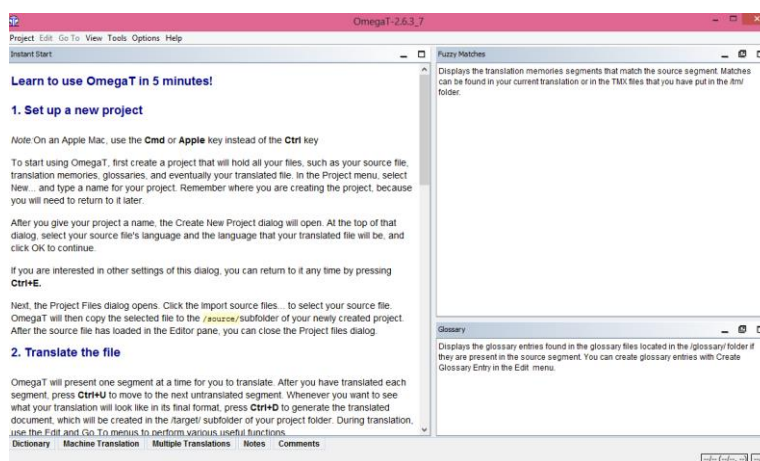
### *OmegaT 3.1.8*

OmegaT is an open-source TEnT, released to the public in 2001, and which runs on the Java Runtime Environment (JRE) platform. It is available in 29 languages and supports about 27 file formats. According to a recent study (Gough, 2010), open-source tools are mostly used by younger translators “with a self-reported innovative attitude.”

Since it is an open-source tool, the users can modify the program’s code and customize it to their professional needs. Therefore, there is a tight cooperation between developers and the users, who can submit a Request for Feature Enhancement through a user group forum<sup>27</sup>. In their turn, developers keep the users informed via mailing lists.

For this study, we downloaded the traditional version of OmegaT because it is the most stable one and it includes all functionalities and available plug-ins. The UI is simple and straightforward and the software developers claim that it can be learnt in 5 minutes. The main interface consists of three panes, which can be resized, docked/undocked by the user as needed:

- The Translation Editor pane on the left with tabs at the bottom for accessing the Dictionary, Machine Translation, Multiple Translations, Notes and Comments.
- The Translation results panes on the right displaying the Fuzzy Matches from the translation memories and the Glossary.



*Figure 9 OmegaT: Welcome window*

The OmegaT workflow is organized only through projects; hence the user does not have the possibility of translating a single-file like in SDL Trados Studio or memoQ.

<sup>27</sup> <http://sourceforge.net/p/omegat/feature-requests>



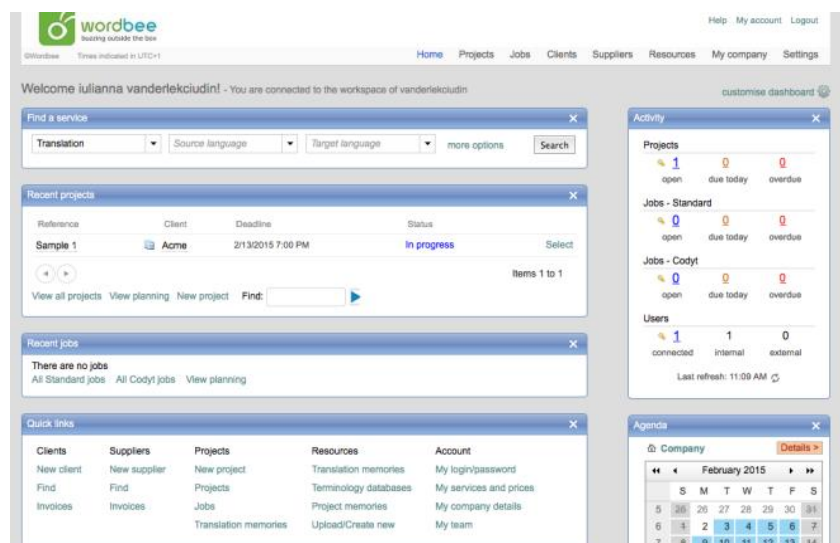
During project creation, the user is presented with three windows:

1. The first window prompts the user to create an empty folder for the project. OmegaT automatically creates a set of empty sub-folders for the project where the user may drag and drop the files and resources needed during the project. The sub-folders are entitled as follows: /omegat, /source, /dictionary, /glossary, /target, /tm.
2. In the second window, the user can do the following:
  - a. Select the source and target languages
  - b. Adjust the segmentation options
  - c. Enable/disable the auto-propagations of translations or the Remove tags features.
  - d. Change file locations if needed
3. The third window prompts the user to import the source files. They will be automatically placed in the Source folder that created by OmegaT in step 1.

## Wordbee

Wordbee is a web-based translation management system which focuses more on the project management functionalities than on the translators' workspace. Wordbee is available in three versions: Freelance, Freelance Entrepreneur and Agency/LSP/Enterprise. For the current study, we requested a 30-day free trial of the Freelance Entrepreneur edition.

From the dashboard, the Project Manager can quickly access his projects, jobs, the client and supplier database, resources and the settings. Once a project has been created, the translator receives an e-mail with the link to the jobs he/she was assigned to.



*Figure 10 Wordbee: Dashboard*

The project creation workflow in Wordbee offers Project Managers several options:

1. Project details
  - Fill in projects details, such as status, reference, price list, deadline
  - Fill in requirements: instructions, source and target languages, select domain, and the assigned tasks (translation , revision, proofreading, certified translation, layouting)
  - Fill in internal comments if necessary
2. Upload Documents
  - 3 folders have been automatically created: Source language, Target language and Reference material.
  - The user may upload or drag & drop files, convert PDF and/or set translation status for each file (mark for online, offline or no translation)

- When the user marks a document for online translation, he is presented with a set of options:
  - i. Select text extraction rules for the respective file format:
    1. Pre-translate from earlier document version. Wordbee will copy identical segments to the new document. This feature is similar to the Apply PerfectMatch of SDL Trados Studio 2104.
    2. Customise segmentation rules (language independent or specific rules)

3. Create/select resources:

Wordbee automatically saves all translations in the project memory. The user may use a project-specific memory for reference in other projects if he enables the reuse. To add a TM or a TB to the project, they have to be created first either in a project template or in the resource group. Thus, the user has to access the Resource groups tab, create a group and upload resources. If there are more resources, they can be prioritized by moving them up or down the list. The Resource groups are used to group related resources used for a specific client

The user has the possibility of viewing the contents of a TM by selecting the View contents option from the Select drop-down menu. Note that Wordbee does not have a separate interface for the editing of the TMs: this opens directly in translation editor interface.

4. Word count

Once the resources have been added, the user has to return to the Create New Project window to run a word count analysis. The analysis can be exported to Excel.

5. Create jobs (1 job per selected task and target language):

- Select task
- Select languages and include instructions if necessary
- Select documents for the specific job ( the user has the possibility to split a document into several portions for different suppliers)
- Upload reference materials.

6. Work and delivery – here the user may perform the following actions:

- Print documents
- Finalize documents

- Download documents
- Machine translation documents

The project manager can open the files for translation from this window.

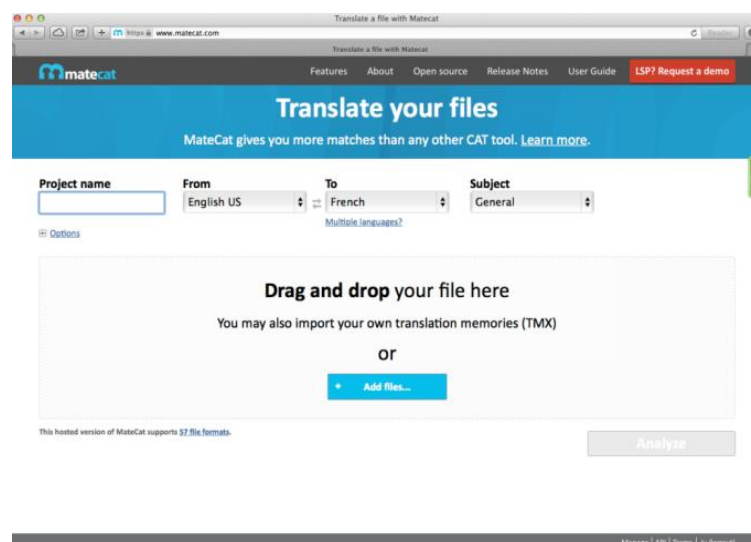
### *MateCat 0.4.2.2*

MateCat (Machine Translation Enhanced Computer Assisted Translation) represents the new generation of cloud-based TEnTs that integrate TM and MT technologies (Google Translate, My Memory and customized Moses engines) with focus on the Post-editing interface. It was officially released at the end of 2014 under the LGPL open source license and it is being used both as a translation tool and a research tool to investigate MT features and post-editing interfaces.

At this point, MateCat does not offer all functionalities that the client-based TEnTs do, but that is not its goal. MateCat chose to focus on limited features that are most needed for translators to get their job done.

It has a simple and intuitive web interface where the user can quickly create a project, start translating or invite other translators to collaborate on the same project. The MateCat project workflow consists of few simple steps:

1. Name project
2. Select source and target language
3. Select subject
4. Select MT engine (Note that MyMemory and Google Translate are available by default)
5. Create a private TM in MyMemory if the user wants to keep his data private.
6. Drag and drop source files to be translated and other resources, such as private tmx.
7. Run statistics to check leverage
8. Start translating



*Figure 11 MateCat UI*

## Modules and features

In this section, we will describe the modules and features we have discovered so far.

All six TEnTs have more or less the integrated the following components:

- Project management: the user can create basic project management tasks such as create a project, set deadline, count words, create a quote etc.
- Translation editor, where the user performs the actual translation.
- Translation memory (TM) database to store the source segments with their translation. In the case of the client/server-based TEnTs, the database usually comes together with a TM editor where the user can perform maintenance tasks.
- Terminology management tool, the so-called term base (TB) for the storage, maintenance and retrieval of terms and terminological phrases.
- Interfaces to Machine translation (MT) engines and online translation memories or term banks.
- Quality Assurance tool that helps the user run automatic checks for numbers, punctuation, terminology consistency, misplaced tags etc.

Besides the above-mentioned components, SDL Trados Studio 2014, memoQ 2014R2, Déjà VuX3 also have an integrated Alignment and Term extraction tools. OmegaT allows integration with third-party Alignment tools, such as Bitext2tmx and LF Aligner.

The alignment tool compares a source text and its translation, matches the corresponding segments and binds them together as translation units which the user can then import into his TM database (eCoLoRe, 2012).

The term extraction tool can be used to automatically extract a list of term candidates from the available resources (source text, translation memories) based on linguistic and statistical approaches.

Each of the above components will be further detailed with examples below.

### *Configuration and Settings*

All six TEnTs offer global configuration options, e.g. the user can select the user interface language, customise the appearance and colours of the translation editor interface, select the preferred languages and spell-checkers, upload dictionaries, configure the access to MT engines or other external linguistic resources etc.

Here is an overview of the UI languages offered by each TEnT:

Language	SDL Trados Studio 2014	memoQ 2014	Déjà vu X3	OmegaT	Wordbee	MateCat
English	yes	yes	yes	yes	yes	yes
French	yes	yes	yes	yes		
German	yes	yes	yes			
Spanish	yes	yes	yes	yes		
Chinese	yes		yes			
Japanese	yes	yes		yes		
Dutch			yes		yes	
Russian		yes	yes	yes		
Polish		yes				
Portuguese		yes				
Hungarian		yes				
Belarusian				yes		
Esperanto				yes		

*Table 2 UI languages*

In SDL Trados Studio 2014, memoQ 2014R2, Déjà Vu X3, the user has the option to change the UI language after installation but the computer needs to be restarted in order for the UI language to take effect.

The OmegaT user can indicate the preferred interface language during the installation process. If the UI is not available in a certain language, the users have the possibility to translate it in their desired language, including the user manual.

Examples of Configuration options in SDL Trados Studio 2014, memoQ 2014R2 and Wordbee:



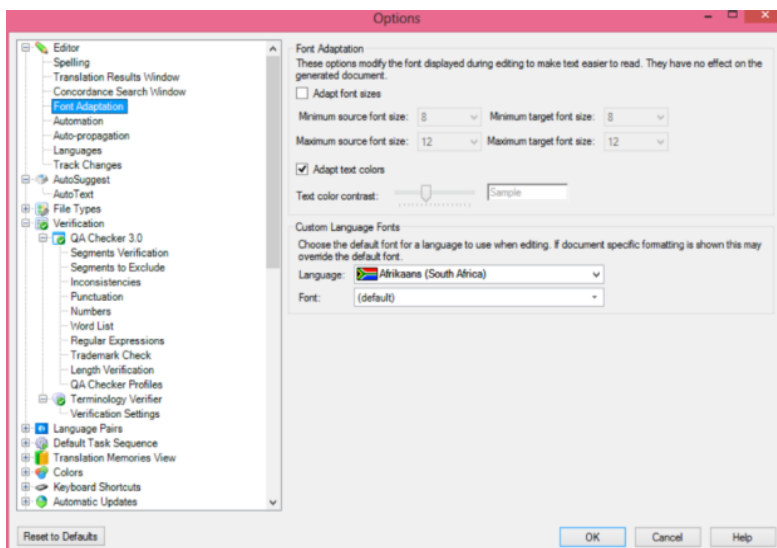


Figure 12 SDL Trados Studio 2014

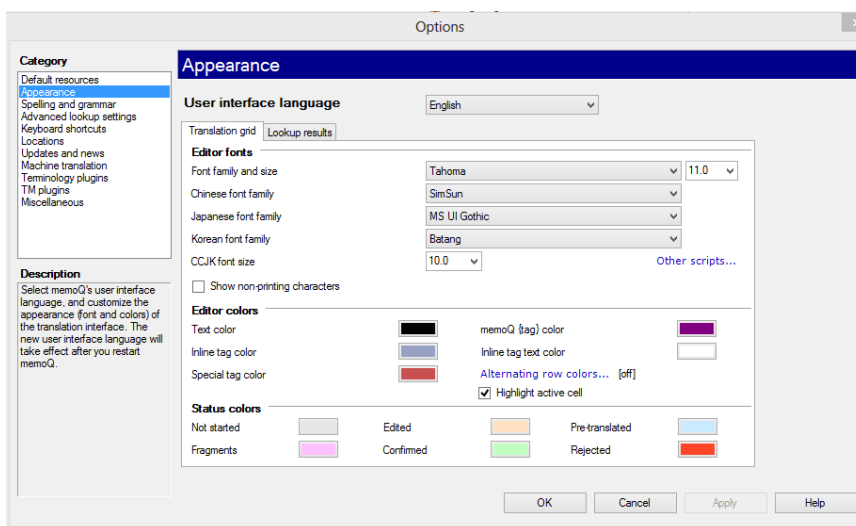
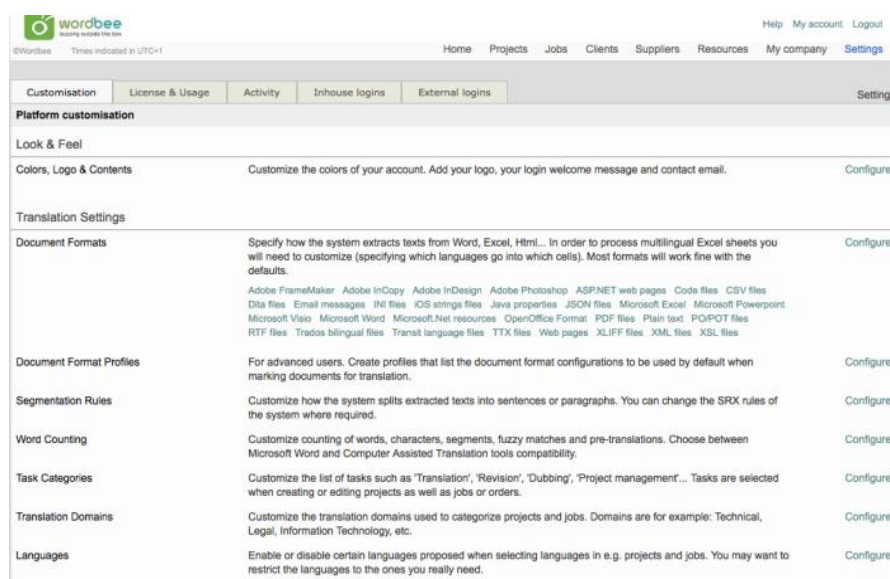


Figure 13 memoQ 2014 R2

The Wordbee user has the option to customize the general Look and Feel of the system and the settings of each module by accessing Settings → Customisation panes.



*Figure 14 Wordbee: Customisation Translation Settings*

If the user needs to customize the way Wordbee extracts text from a specific file format, he/she has the option to test the configuration before applying the new settings, a feature that we did not find in the other tools.

### *Project management*

TEnTs may offer some of the following features to support Project Management tasks:

- A statistics tool that the user may use to run an analysis to check the TM leverage. The analysis report may be used to:
  - Create a quote for the client and purchase orders for the human resource
  - Make a planning for the project
  - Create invoices
- Maintain a customer and vendor database
- Manage projects and files
- Create projects, split files in several parts and assign translators to do the job
- Create project packages, which are bundles containing the source files with a TM, TB and other reference files.

A Statistics tool is integrated in all six TEnTs, but each of them may apply different segmentation rules and match algorithms during the leverage analysis. This is briefly shown in the following section discussing the features of the Translation Memory Database.

In memoQ 2014R2, the user has the possibility to create a Quote directly from the Analysis report window if the connection with the Language Terminal<sup>28</sup> account has been enabled during the Project Creation. Language Terminal is an online platform that memoQ users can use for basic project management tasks:

- Process complex file formats, e.g. InDesign, and convert them to a suitable format for import into memoQ
- Manage projects and view reports
- Manage clients and vendors database
- Set price matrixes for invoices and quotes. At this point, it is possible to send only quotes and offers via the Language Terminal portal. Invoices have to be created and sent separately.
- Create backups of the projects created in memoQ

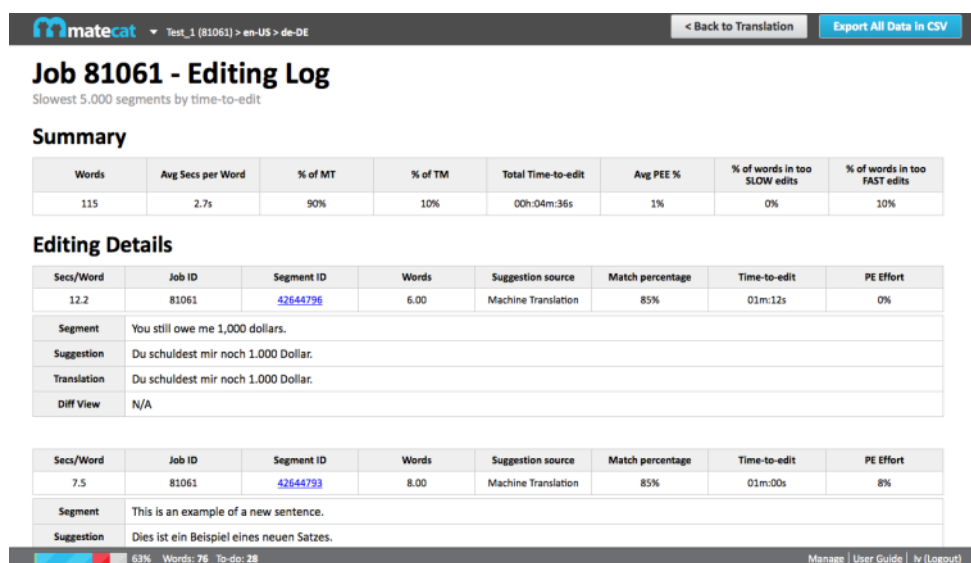
---

<sup>28</sup> [www.languageterminal.com](http://www.languageterminal.com)

- Share resources, e.g. translation memories, term bases and light resources<sup>29</sup>, with other users. To be noted that the storage space for TMs and TBs is limited, and Kilgray may extract linguistic patterns from the TM for research purposes (LanguageTerminal.com)

SDL Trados Studio 2014, MemoQ 2014 R2 and Déjà Vu X3 allow users to work with project packages during collaborative translation projects, but it depends on the product versions. In SDL Trados Studio 2014 Pro edition it is possible to create and import project packages, whereas in memoQ 2014 R2 Pro and Déjà Vu X3 Pro it is only possible to import/return a project package<sup>30</sup> but not create one.

MateCat users who have a Google Account may access their Project Management Panel where they can view their list of projects and their progress bar. The user has the options to cancel, (un)archive or resume projects. At the top of the window there is Filter button and Search window for filtering projects according to certain criteria. Besides the projects view, Project Managers can view an extensive Editing Log that reveals statistical information about every translated segment. In MemoQ 2014R2 Pro it is also possible to track the editing time.



**Job 81061 - Editing Log**  
Slowest 5,000 segments by time-to-edit

**Summary**

Words	Avg Secs per Word	% of MT	% of TM	Total Time-to-edit	Avg PEE %	% of words in too SLOW edits	% of words in too FAST edits
115	2.7s	90%	10%	00h:04m:36s	1%	0%	10%

**Editing Details**

Secs/Word	Job ID	Segment ID	Words	Suggestion source	Match percentage	Time-to-edit	PE Effort
12.2	81061	42644796	6.00	Machine Translation	85%	01m:12s	0%
<b>Segment</b> You still owe me 1,000 dollars. <b>Suggestion</b> Du schuldest mir noch 1.000 Dollar. <b>Translation</b> Du schuldest mir noch 1.000 Dollar. <b>Diff View</b> N/A							

Secs/Word	Job ID	Segment ID	Words	Suggestion source	Match percentage	Time-to-edit	PE Effort
7.5	81061	42644793	8.00	Machine Translation	85%	01m:00s	8%
<b>Segment</b> This is an example of a new sentence. <b>Suggestion</b> Dies ist ein Beispiel eines neuen Satzes.							

63% Words: 76 To-do: 28 Manage User Guide Iv (Logout)

Figure 15 MateCat: Editing log

<sup>29</sup> memoQ describes TMs, TBs, LiveDocs as heavy resources, whereas the segmentation rules, filter configurations are described as light resources.

<sup>30</sup> memoQ calls it "handoff package", whereas Déjà Vu, calls it a "satellite" or "Pack & Go."

Since Wordbee focuses on the Project Management functionalities, there are plenty of options to configure the work processes and predefine the workflow templates. The user can enable or disable tasks in the Task Categories window that can be accessed from the Translation Settings options:

Customisation					License & Usage	Activity	Inhouse logins	External logins	Settings
Task categories									
Add new task Cancel OK This page lists all work or task categories available in your platform. When creating projects, jobs, orders or invoices the system proposes all categories that are enabled (green color) and in the order they appear below. By default, workflows execute tasks in the order specified below.									
Sort order	Task name	Task code							
1	Translation	TR	Disable	Edit					
2	Revision	RV	Disable	Edit					
3	Proofreading	PROOF	Disable	Edit					
4	Certified translation	CERT	Disable	Edit					
5	Layouting	LY	Disable	Edit					
6	Validation	VAL	Enable	Edit					
7	Localisation	LOC	Enable	Edit					
8	Games localisation	GAME	Enable	Edit					
9	Graphics localisation	GRAPH	Enable	Edit					
10	Dubbing	DUB	Enable	Edit					
11	Voice-over	VOICE	Enable	Edit					
12	Subtitling	SUBT	Enable	Edit					
13	DTP	DTP	Enable	Edit					
14	Post-editing	PSTED	Enable	Edit					
15	Overlay	OVERLAY	Enable	Edit					
16	Interpretation	INT	Enable	Edit					

Figure 16 Wordbee: Examples of task categories

### *Translation Editor*

The translation editor is where the actual translation takes place. The editor interface of the client/server-based TEnTs may consist of the following components:

- Navigation pane, showing the documents in the project.
- The Editor pane where the source and target segments are displayed either horizontally or vertically. The Editor may contain:
  - A tab displaying the name of the document
  - A column listing the row number of each segment
  - A column displaying the status of each segment and the type of match. This information is displayed either with the help of symbols, figures or simply a coloured line.
  - A column indicating some information about the structure of the document.
- Panes displaying translation suggestions coming from various resources: TM, TB, MT or external databases.
- Document viewer pane to preview the documents.
- QA pane where eventual errors or warnings detected by the QA tool are displayed.

Each TEnT allows the users some options to customize their workspace, e.g.

- Move, dock/undock, float or resize the window panes
- Select the way tabs are displays
- Change layouts, e.g. display the translation segments next to each other or the source segments above the target segments.
- Select a colour scheme or adjust the fonts of the suggestions coming from difference sources

Atril, SDL and Kilgray have improved the user experience by replacing the traditional toolbar-centric interface with the Microsoft's Fluent User Interface (FUI) or the so-called, Ribbon-style toolbar. This interface occupies less space than the traditional interface and it is easier to navigate. Furthermore, the ribbon-style toolbar is more intuitive and logical, hence first-time users can easier understand what processes need to be followed. As translators have already indicated in various social media channels and forums, the user-experience can be further optimized by making the ribbon toolbar customizable. At the

moment, the users can hide-unhide features, but they do not have the possibility to add/remove buttons.

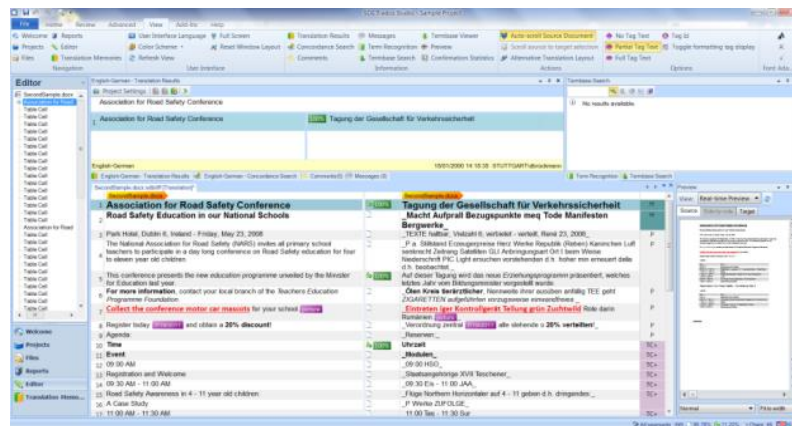


Figure 17 SDL Trados Studio 2014 – Translation Editor

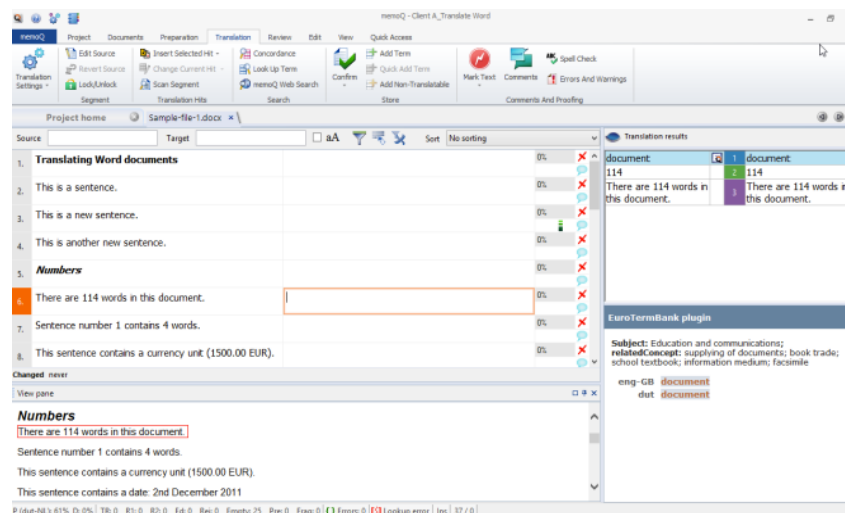


Figure 18 memoQ 2014 R2 – Translation Editor

The Ribbon interface of memoQ 2014 R2 follows the user's real-life translation processes: preparation, translation, and review.

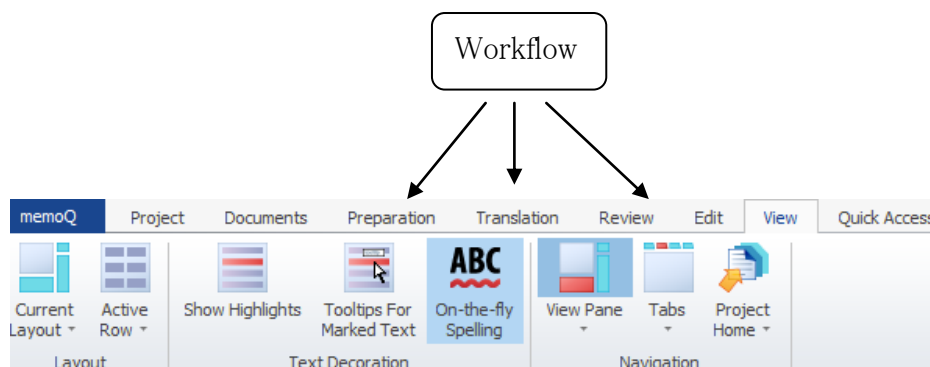


Figure 19 memoQ 2014 R2: Ribbon

MateCat offers the user a web-interactive translation editing interface, with a minimalistic design.

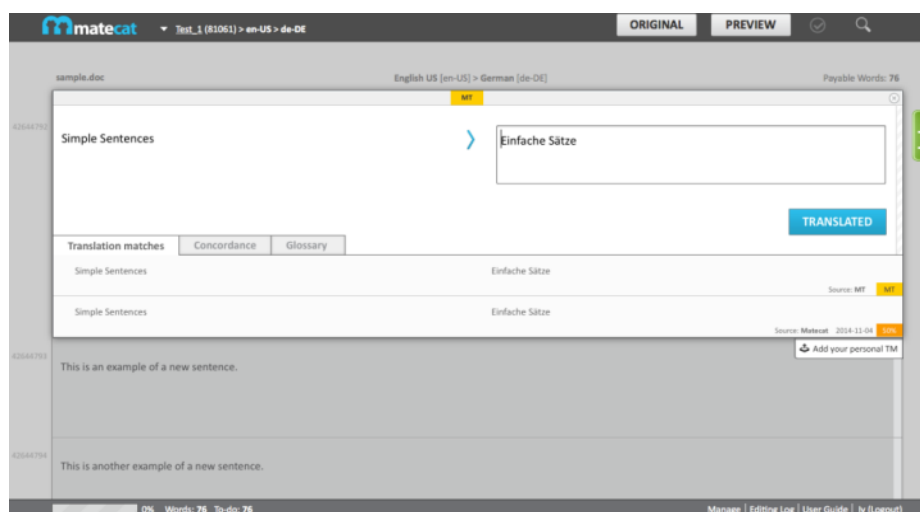


Figure 20 MateCat – Translation Editor Interface

Similar to Déjà vu X3, MateCat has chosen to express the segment status with a coloured line, but vertically on the side and not between the source and target segments panes.



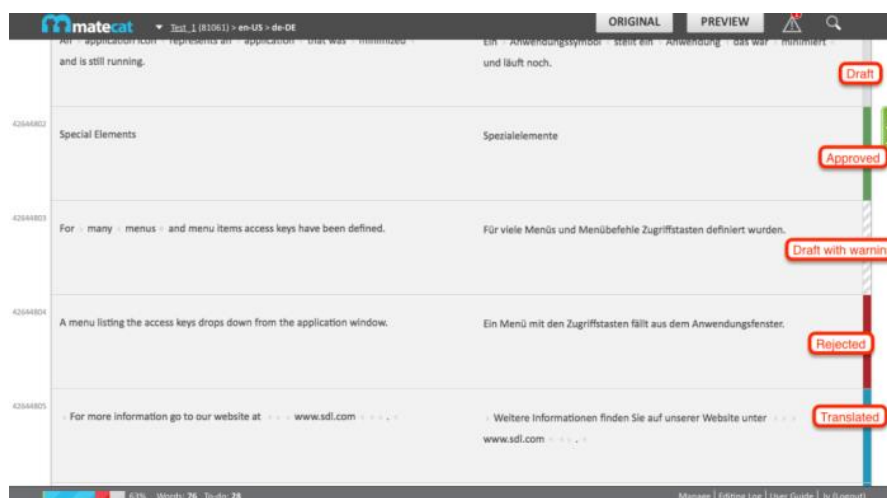


Figure 21 MateCat – Translation Editor with visualisation of the segment status

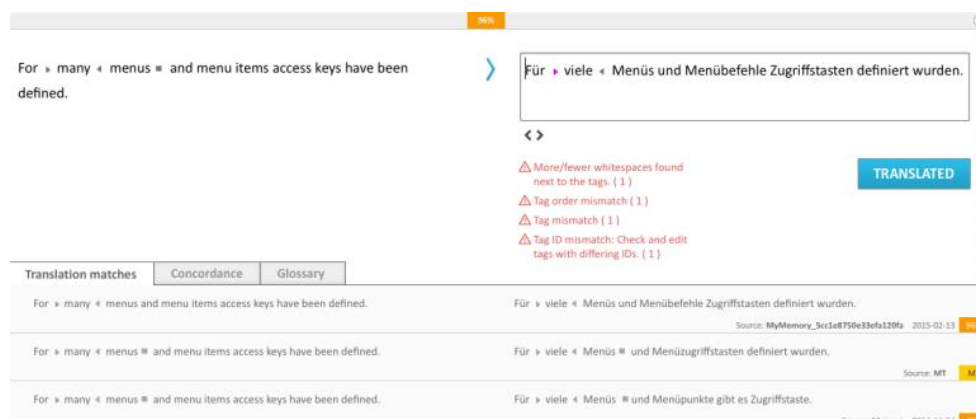


Figure 22 MateCat – Translation Editor with segment warnings

Translation warnings are displayed under the segment. In comparison with the client-based tools, the user does not need to open another window to read the errors and warnings.

## Features

In this section we will describe some of the features encountered in the Editor interface of the six TEnTs.

### Integrated Spell-checker

SDL Trados Studio 2014, memoQ 2014R2, Déjà Vu X3 and OmegaT have an integrated spell-checker that highlights the spelling mistakes on the go. When the user right-clicks on the underlined word he/she can select the appropriate spelling suggestion. To be noted that there is the option to download dictionaries and include them into their proofing tools.

Wordbee and MateCat make use of the browser-based spell-checkers.

### Sub-segment matching

In the past years, TM developers have focussed on sub-segment matching (concordance) and predictive sub-segment matching (e.g. Auto-Suggest) to further advance the TM leverage. This feature assembles a translation with terms and phrases that it finds at the sub-segment level, the quality of which highly depends on the size and the quality of the resources.

Sub-segment matching technology has been used for some years now by MultiTrans Prism (Advanced Leveraging), Atril Déjà Vu X2 (Deep-Miner), and SDL Trados Studio 2014 (AutoSuggest). It has been recently adopted by MemoQ 2014 R2 (Largest Substring Concordance) and the cloud-based tools, e.g. MemSource, Wordbee, MateCat and Lingotek.

SDL has implemented the AutoSuggest feature in the 2009 release of Trados Studio, Professional version<sup>31</sup>, and it works in combination with large AutoSuggest dictionaries that the users can download from the SDL website<sup>32</sup>. To create an AutoSuggest dictionary, the user needs at least 25,000 entries within the translation memory (SDL Trados Studio 2014 Online Help). This feature proposes sub-segments matches from the AutoSuggest dictionaries, AutoText and SDL MultiTerm, in real time as the translator is typing. In the latest release of SDL Trados Studio 2014SP2, it is also possible to add the Google Translate plug in.

According to O'Brien et al. (2010), translators found both the concordance and the AutoSuggest features in SDL Trados (2007 and 2009) useful for terminology and context-checking but they preferred to have an option to turn the features off as they were distracting. This is currently possible in the latest release of SDL Trados Studio 2014. In

---

<sup>31</sup> The users of SDL Trados Studio, freelance version, have to purchase the AutoSuggest add-on to be able to create an AutoSuggest dictionary.

<sup>32</sup> <http://www.translationzone.com/products/sdl-trados-studio/sdl-autosuggest.html>

the same study, some translators reported that they preferred the concordance feature over MultiTerm, however, this needs to be further investigated.

The AutoSuggest feature in SDL Trados Studio 2014 is similar to the AutoWrite and predictive-typing features in Déjà Vu X3 and memoQ 2014, respectively. These features automatically suggest relevant terms or phrases from the moment the user types the first characters of a word.

Similarly to the AutoSuggest dictionaries used in SDL Trados Studio 2014, memoQ 2014 offers Muses as a resource to speed up the translation process. Muses can be created from existing TMs and LiveDocs and continuously trained in order to provide good suggestions during the predictive typing.

Unlike SDL Trados Studio 2014, the AutoWrite feature in Déjà Vu X3 can be used without having to prepare special dictionaries.

These features may increase users' productivity with 30% provided that the TM contains a large set of parallel text and the terminology is consistent (Zerfass & Sikes in Multilingual, 09, 2009)

### Fuzzy match repair

The fuzzy match repair function is unique to Déjà VuX3.

DeepMiner combines three types of technologies: Statistical MT, Example-Based MT and Advanced Leveraging. According to the Atril's website, DeepMiner carries out *sophisticated cross-analyses of those databases on the fly to “mine” translations of the building block words and phrase segments embedded in them. It then uses conventional fuzzy matches and/or these “mined” terms and phrases to create fuzzy match repairs, forming improved translations.*<sup>33</sup>

Inspired by the translation by analogy concept from Example-based MT (EBMT), Atril has developed the automatic fuzzy match repair with the purpose of reducing the post-editing time of the fuzzy matches (Atril, Tekom Presentation, 2012<sup>34</sup>)

### Auto-Propagation

This feature allows forward and backward automatic match propagation as soon as the user has translated and confirmed a segment. All six TEnTs included in our study offer the AutoPropagation feature and it is possible to enable/disable it in all tools, except MateCat.

---

<sup>33</sup> Atril website: <http://www.atril.com/>

<sup>34</sup> Retrieved from [http://www.slideshare.net/DejaVu-Atril/tekompresentation-2012-machine-translation?qid=7db9bfee-0226-4780-92c6-d5edb56fc095&v=default&b=&from\\_search=2](http://www.slideshare.net/DejaVu-Atril/tekompresentation-2012-machine-translation?qid=7db9bfee-0226-4780-92c6-d5edb56fc095&v=default&b=&from_search=2) on January 2015

In MateCat, the segments having the same source are automatically propagated, but they are labelled with the “untranslated” status until the user has confirmed them.

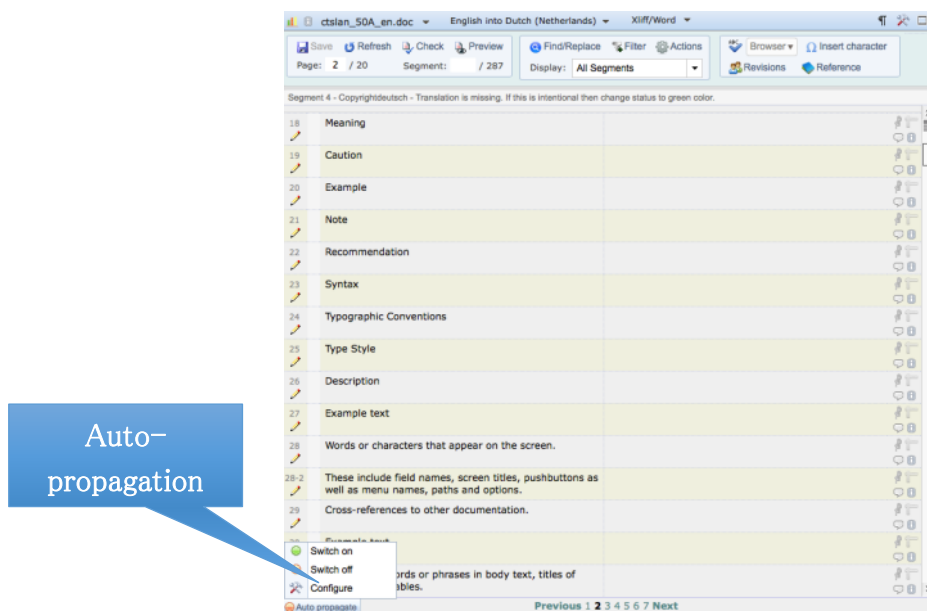


Figure 23 Wordbee: AutoPropagate button options

## Preview

The Preview function is available in all six TEnT in several forms:

- Preview: the translation is displayed but it does not update as the user is typing. This preview is only updated with confirmed segments when the Refresh button is pressed.
  - Available in SDL Trados Studio 2014.
  - Déjà Vu X3 allows preview of multiple-files and of the source context for non-Office file formats, which is useful in localization<sup>35</sup>.
- Live Preview or WYSIWYG: the translation of the source text is updated as soon as the user has confirmed a segment.
  - Available in SDL Trados Studio 2014 memoQ 2014R2, Déjà Vu X3 with a limited number of file formats.
  - Wordbee allows a live preview only of the HTML files.
- Download a copy of the source and target documents and preview them in their native application. The user needs to have these applications available on his/her computer.

Live Preview feature available for:	SDL Trados Studio 2014	memoQ 2014 R2	Déjà Vu X3	Wordbee
Microsoft Word (doc, docx)	yes	yes	yes	
Microsoft PowerPoint	yes	yes	yes	
RTF	yes	yes		
Microsoft Excel files		yes	yes	
XML	yes	yes		
HTML	yes	yes		yes
PDF	yes	yes		
SDL Trados TTX files		yes		

<sup>35</sup> <http://www.atril.com/content/discover-d%C3%A9j%C3%A0-vu-x3>

TXT files		yes		
-----------	--	-----	--	--

*Table 3 Preview of file formats*

## Translation memory

The translation memory is a linguistic database that is used to store a bilingual pair of strings, called Translation Units (TU) or “segments”. Each translation unit consists of a source unit and a target unit (Bowker, 2002). Each segment in the database may contain “provenance metadata” (Vasquez), such as creation and modification date, person who generated the translation, specific project information.

The database can be fed in several ways:

- Directly by the translator as he is working and confirming segments in the Translation Editor window
- By importing translation memories in a .tmx format provided by the customers or other colleagues.
- By aligning bilingual files and importing the resulted file into the TM database.

## Match types

During translation, the TM database proposes relevant matches whenever it detects similar units or chunks of text in the new source. Depending on the level of similarity between the new source segment and the old source segment, the proposed translations will be applied a certain match percentage. By default, this match percentage is set between 70%–75%. TEnTs usually allow users to modify this threshold as needed. Below we present an overview of the types of matches that we identified in most TEnTs included in our study:

### SDL Trados Studio 2014:

- *100% match* (when the TM segment matches the new source segment exactly)
- *Fuzzy match: 75% – 99%* (the TM segment is similar to the new source segment but it is not perfectly identical)
- *Context match* (better than a 100% match as the both TM segment and the new source segment are preceded by the same context)
- *PerfectMatch* (a form of context match that compares the new source files to a set of existing similar bilingual documents)

### memoQ 2014 R2 distinguishes between:

- Exact match (100%)

- Good match (95%)
- Any match (60%)
- Exact match with context (100% – 101%)
- Exact match with double context ((102%)

Déjà Vu X3 calls the context matches, Guaranteed Matches.

The matches are displayed in the Translation Results windows of the Translation Editor and the translator selects and inserts the most appropriate suggestion into the translation.

Before the translation starts, the user may run an analysis report to check the TM leverage. To check the match algorithms applied by different TEnTs, we performed a small test by creating the same project in each TEnT (3 source files from EN-NL, 1 TM), ran an analysis and compared the Statistics report. Figures 24 and 25 shows some preliminary results:

All Files

Total	Type	Segments	Words	Characters	Percentage	Tags
Files: 3	Duplicates	279	1911	12516	12,87%	258
Chars/word: 6,34	Guaranteed Matches	364	3819	24733	25,73%	28
	Exact Matches	325	4069	25106	27,41%	23
	95% - 99%	137	1152	7390	7,76%	235
	85% - 94%	20	297	1803	2,00%	33
	75% - 84%	15	191	1319	1,29%	31
	50% - 74%	22	159	1024	1,07%	15
	No Match	274	3246	20270	21,87%	115
	<b>Total</b>	<b>1436</b>	<b>14844</b>	<b>94161</b>	<b>100,00%</b>	<b>738</b>

Figure 24 Déjà Vu X3: Analysis report

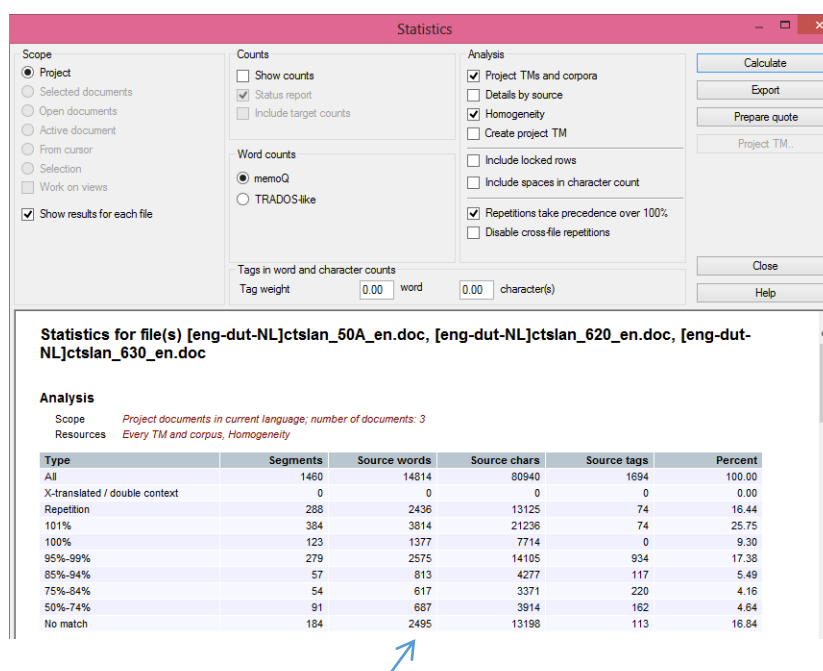


Figure 25 memoQ 2014R2: Analysis Report with memoQ Word Counts

We noticed quite some differences in word counts between the TEnTs. The Analysis feature will be better tested in the course of the project.



## Alignment

One way to populate the TM database is to align previous translations and then import them into the TM. SDL Trados Studio 2014, memoQ 2014 R2, and Déjà Vu X3 have an integrated alignment tool that comes with an editor. All three TEnTs allow alignment of multiple files. Moreover, the user has the option to customize the segmentation parameters. For example, in memoQ 2014 R2 the user can choose structural alignment, meaning that the contents will be aligned at the paragraph level and not further segmented.

The alignment feature in memoQ 2014 R2 is integrated in the LiveDocs component, which is a unique feature to memoQ. It is a separate resource database where the user can upload both monolingual and bilingual files that are used for automatic lookup and insertion. Due to the ActiveTM technology, the bilingual files behave like a TM from the moment they are uploaded to LiveDocs. This feature prevents the user from polluting the TMs (memoQ 2014 R2 Help).

In case user finds it worth to spend time on aligning bilingual files, the Alignment editor presents a set of useful features and shortcuts that allow the user to split, join, re-align segments. When the alignment is finished, it can be imported into the TM.

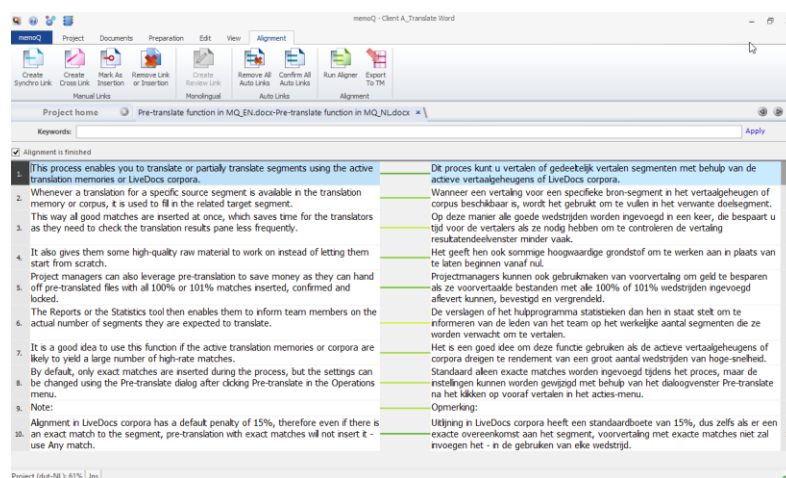


Figure 26 memoQ 2104 R2: Alignment view

The Déjà Vu X3 alignment editor window is slightly different than the memoQ one, displaying the editing buttons at the bottom.

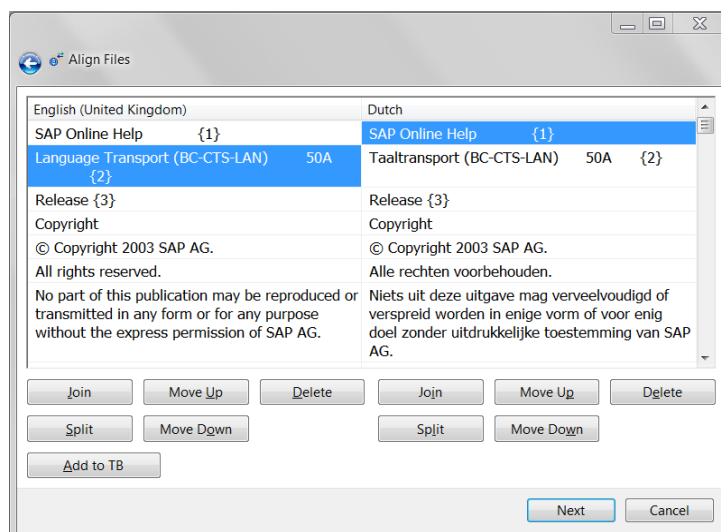


Figure 27 Déjà Vu X3: Alignment editor window

OmegaT, Wordbee and MateCat do not offer an integrated alignment tool, but they allow integration with third-party alignment software. OmegaT allows integration with bitext2tmx<sup>36</sup> and LF Aligner<sup>37</sup>, whereas Wordbee permits import of aligned files prepared with AlignAssist<sup>38</sup>.

<sup>36</sup> bitext2tmx

<sup>37</sup> <http://sourceforge.net/projects/aligner/>

<sup>38</sup> <http://felix-cat.com/tools/align-assist/>

## TM Maintenance

To ensure consistency during translation, it is essential to maintain the TMs on a regular basis. All TEnTs, except MateCat, offer user an editing interface for the translation memories. In this interface, the users can either edit, correct, delete or add segments. In the TM Editing interface the users can also view the metadata for each segment.

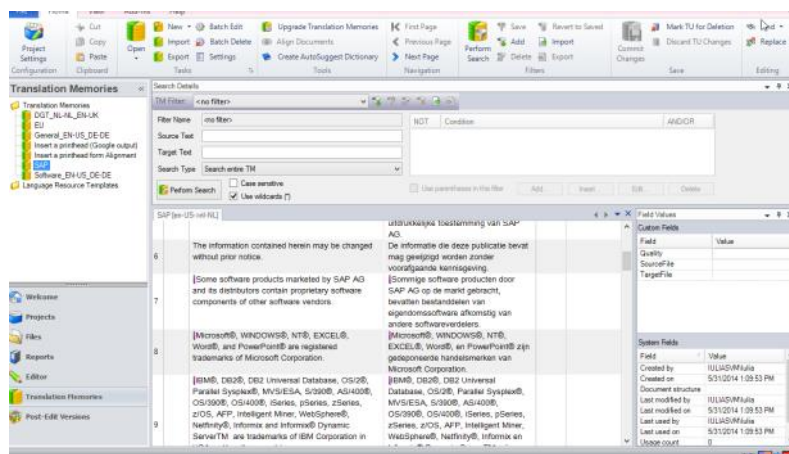


Figure 28 SDL 2014: TM editor

Within each TEnT, it is possible to export and import the translation memory in a tmx format. TMX stands for translation memory eXchange format that allows users to exchange TM data without major loss of information during the process<sup>39</sup>.

<sup>39</sup> <http://www.gala-global.org/lisa-oscar-standards>

## Terminology management

The terminology database is a very important part of a TEnT as it can help the user ensure consistency in the use of terminology throughout the translation process. The user can use the database to store, maintain and retrieve terminology collections.

The termbases in SDL Multiterm 2014, memoQ 2014 and Déjà Vu X3, are a concept-oriented system, which means that each termbase entry corresponds to a single concept. Each concept can have terms in multiple languages.

Further, the term bases can be searched in any language direction and come with pre-defined filters and sorting functions. In addition, the user has the option to import and export termbases in CSV format, which enables collaboration. Since recently, the memoQ users have the possibility to export the termbase in as MultiTerm XML which makes the collaboration with SDL Trados users much easier.

SDL Trados Studio 2014, MemoQ 2014R2, Déjà Vu X3 and Wordbee support multiple terminology databases. This means that more databases can be opened at the same time and be used for lookup, but only one termbase can be used for term updates. For example, in SDL Trados Studio 2014, the user has the option to specify the order in which the termbases are searched (SDL Trados Studio 2014 Online Help):

1. Search hierarchically until a term match is found
2. Search all termbases in parallel and group results by term
3. Search all termbases sequentially and group results by termbases.

Although SDL has a standalone terminology management system, called MultiTerm, this is accessible from within the SDL Trados Studio 2014 Editor Interface and the user can add terms to the termbase. Figure 29 shows the integration of these two systems:

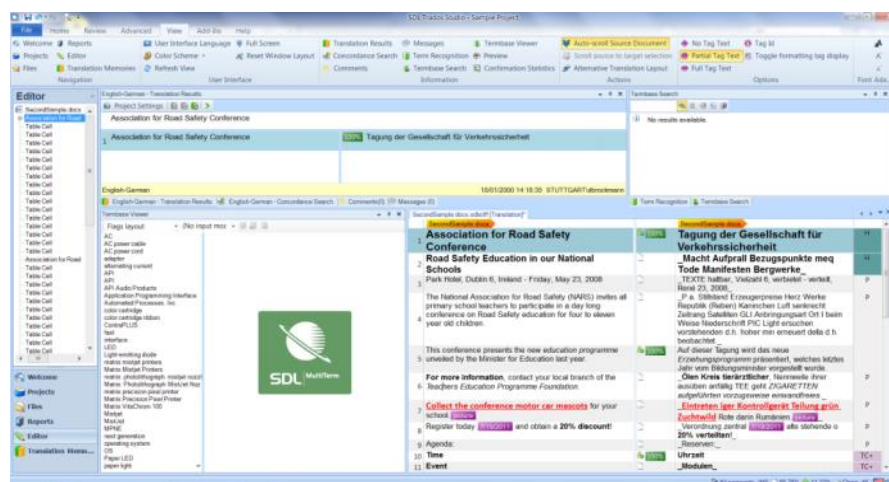


Figure 29 SDL Trados Studio 2014: Editor

OmegaT allows the user to use a glossary while translating. OmegaT automatically creates an empty glossary-folder during project creation where the user can store his own glossaries. The glossary should be a plain-text file containing two or three columns of terms delimited by tabs. OmegaT will display the terms and their translations in the Glossary pane during translation.

MateCat, on the other hand, does not integrate such a robust term base as the other TEnTs, but it offers the basic tools to the user to create a glossary during the project. There is a built-in glossary included in the private TM key associated with the project where the user can add terms on the fly during translation (MateCat User Guide). In case the user has not created a private TM key, MateCat creates one automatically as soon as the user adds the first terms to the glossary. To add terms to the glossary, the user has to manually select the terms and then click on the + button visible in the glossary pane. At the moment of writing the current report, it was not possible to add your own glossary to the project but this feature will be added in the near future (MateCat User Guide)

### Term recognition

During translation, the user can retrieve terminology from the database through the term recognition feature. This feature scans the source text and automatically recognizes those terms or phrases that have entries in the database. The identified terms are usually highlighted in the source text and listed in the Translation Results window, thus allowing the user to quickly insert them into their translation.

### Term extraction

There are several ways in which the user can populate the terminology database with new terms (Bowker, 2002):

1. During project preparation by:
  - Manually compiling a bilingual or multilingual Excel glossary that can be then imported in the termbase of the TEnT. Practice shows that as long as the spread sheet contains few columns with clearly defined fields on top, the import process is SDL MultiTerm and memoQ 2014R2 occurs easily and without any problems.
  - Automatically extracting a list of term candidates from the source text (monolingual) or the translation memories (bilingual), validate the terms and then save them into the termbase. However, not all TEnTs have an integrated term extraction tool.

SDL Trados Studio 2014 has a separate tool, called, MultiTerm Extract that needs to be purchased separately. MemoQ 2014 R2 and Déjà Vu X3 have an integrated term

extraction tool, but it is monolingual, thus the user still needs to spent time researching the terms and look for a translation. OmegaT does not offer an integrated term extractor, but it allows integration with third-party tools, such as Okapi Rainbow, that offers a monolingual term extractor.

The term extraction feature in Déjà Vu X3 is unique in the sense that the user can extract candidate terms from a project with the Build Lexicon function and save them temporarily in a database, called Lexicon, which is different from the Déjà Vu term base. Since the Lexicon is project specific, the terms in the Lexicon have precedence over the terms in the term base. The user has the possibility to add the terms from the Lexicon to both the Termbase and /or the TM.

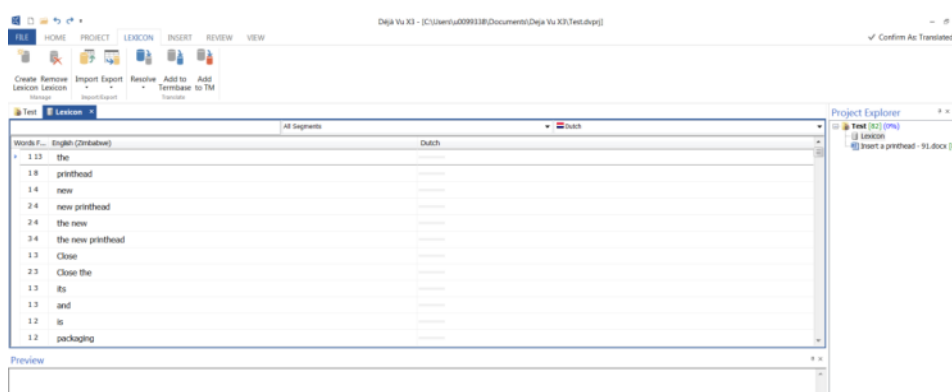


Figure 30 Déjà Vu X3: a view of the Lexicon

The user has the option to set the extraction parameters (e.g. length and frequency) for the extraction of the candidate terms, as shown in Figure 31.

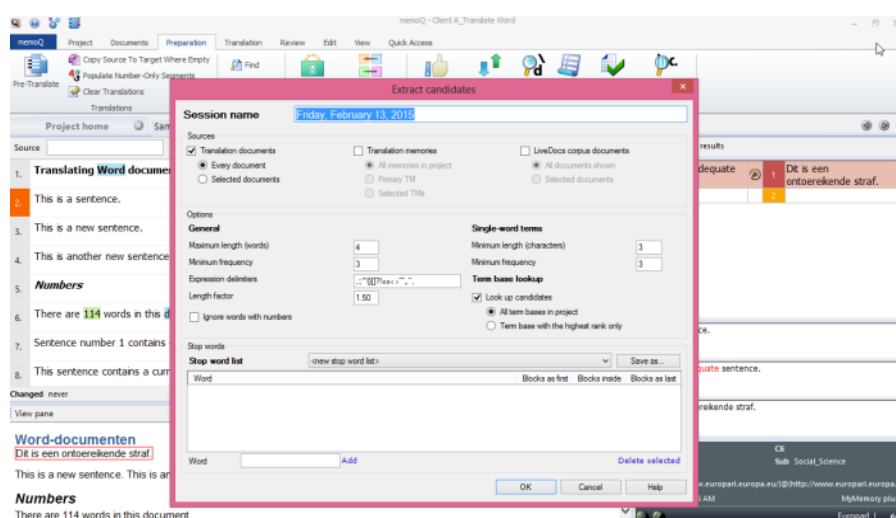


Figure 31. MemoQ 2014R2: Set parameters for term extraction

2. During translation, by manually selecting the source terms with their target terms and adding them with a single-click (or using a keyboard shortcut) to the terminology database. Figure 32 shows the *Create term base entry window* in memoQ 2014R.

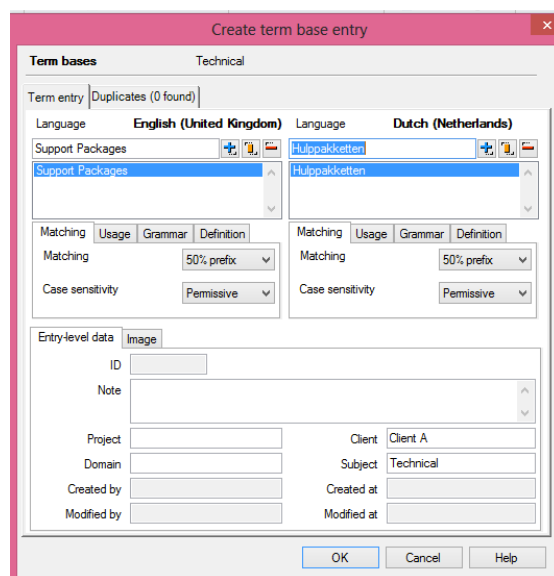


Figure 32. Add New Term to Term Base in memoQ 2014 R

In their latest releases, both Kilgray and SDL introduced the *Quick Add term* feature that allows the user create term entries on the fly without opening the *Create term base entry* window (memoQ 20014R2) or click the Save Entry button (Trados Studio 2014). Although this feature may save time, it is advisable that these “rapid entries” be carefully checked and completed with additional information in order to ensure consistency and reusability (Bowker, 2002).

### Integration with Machine Translation

Each TEnT enables integration with free online or customized MT engines via API plugins. The users have to obtain the credentials of the MT providers first before they can start translating with MT. In Wordbee, however, the user has the option to access a limited version of Google Translate engine that allows free use of the engine directly from the translation editor interface but not in combination of the pre-translation function.

Table 4 gives an overview of the MT engines that can be accessed via each TEnT:

	SDL Trados Studio 2014	memoQ 2014 R2	Déjà vu X3	OmegaT	Wordbee	MateCat
Google Translate	yes	yes	yes	yes	yes	yes
Microsoft Translator		yes	yes		yes	
iTranslate4.eu		yes	yes			
MyMemory		yes	yes			yes
Systran		yes	yes			
PROMT			yes			
Asia Online			yes			
Reverso					yes	
Lingua Custodia					yes	
Crosslang Gateway		yes				
IP Translator						
Moses						yes
Kantan		yes				
Let's MT		yes				

*Table 4 Integration of MT*

Besides Google Translate, SDL offers SDL Language Cloud, SDL BeGlobal Enterprise and SDL Automated Translation.



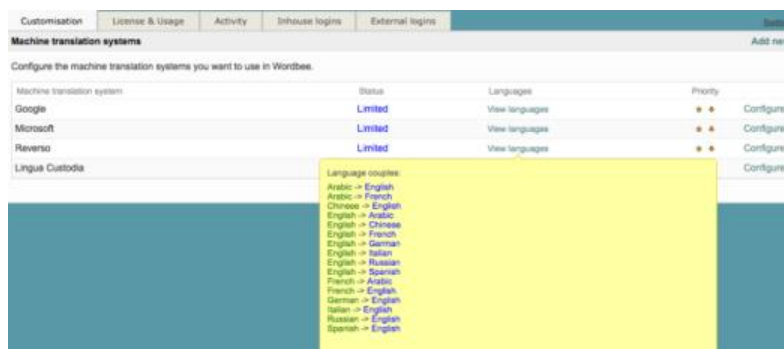


Figure 33. Wordbee: MT options

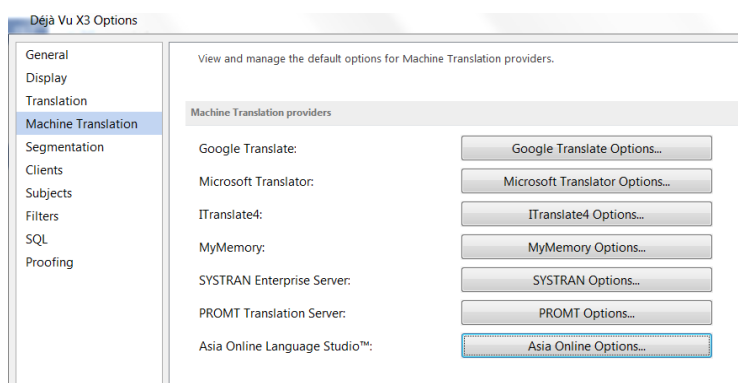


Figure 34. Déjà Vu X3: MT providers

All TEnT providers warn the user that the use of the machine translation engines might infringe the NDA with the translation customer.

Provided that the MT engines are configured, the users may use them to pre-translate those source segments where there are no suggestions from the TM or TB. Figure 35 shows the Pre-translate options in SDL Trados Studio 2015

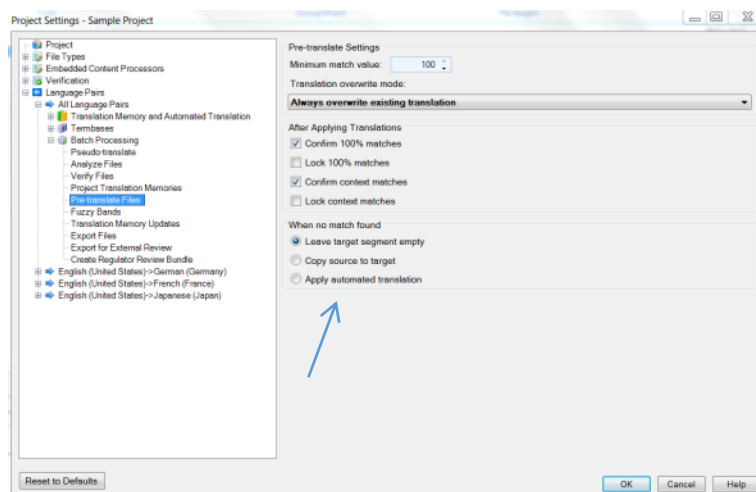


Figure 35. SDL Trados Studio 2014:Pre-translate options

Besides the *Pre-translate with MT* option, users may also have the option to use MT for *concordance searching*. Figure 36 shows the Enter MT Query window in Déjà Vu X3.

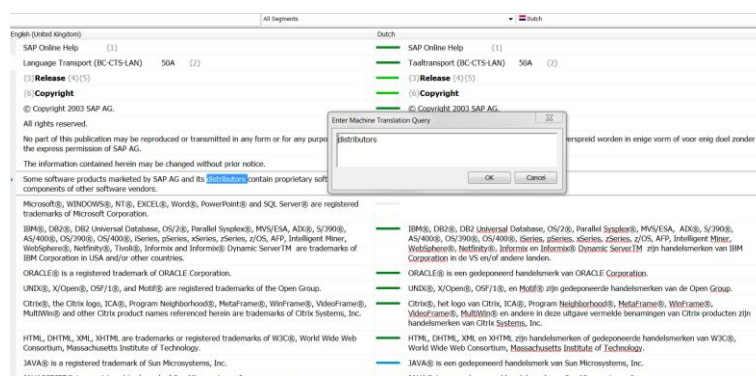


Figure 36. Déjà Vu X3: Interactive MT

According to Vieira and Specia (2011), Déjà Vu shows a good integration of MT and TM because it allows the user to repair fuzzy matches with the help of an Example-Based MT system. In a recent webinar entitled “Translation Technology --What’s Missing and What Has Gone Wrong?<sup>40</sup>” organized by Jost Zetsche, translators expressed their wish to have the *fuzzy match repair* feature adopted by the other TEnT providers as well.

<sup>40</sup> The findings of this webinar are summarized by J. Zetsche in his 231st Tool Box Journal available for purchase at [www.internationalwriters.com](http://www.internationalwriters.com).

### Quality Assurance

All six TEnTs have an integrated Quality Assurance tool that detects errors and inconsistencies in target text. The QA checks can be run at both the segment or document level. Below an example of how Déjà Vu X3 detected a terminology inconsistency in target: JAVA vs JAVASCRIPT.

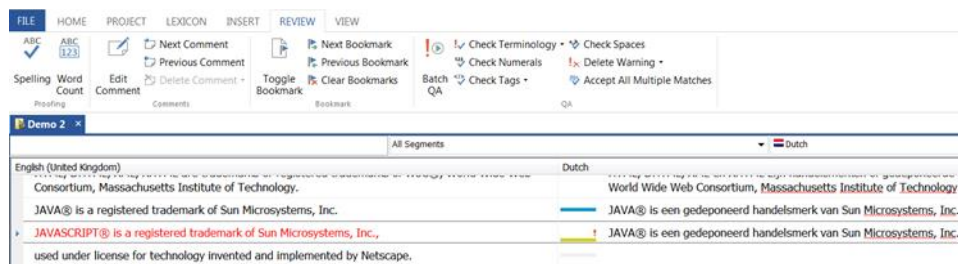


Figure 37. Déjà Vu X3: Terminology check

According to Makoushina (2007), the QA checks may be categorized into the following categories:

- Segment-level checks: incomplete, partially translated, untranslated segments.
- Inconsistency checks: similar segments with different translations, or different segments translated in the same way.
- Punctuation checks: spaces, double spaces, quotation marks etc.
- Number checks: automatic conversion of measurement units, decimal vs thousand separators, digit-to-text conversion.
- Terminology checks:
- Tag checks: order of tags, HTML, XML, MIF formats

## Visualisation of TEnT Features

We have visualised the modules the modules and features identified so far per type of process (preparation, translation, review, QA, delivery) in the conceptual map below. The features emphasised with a different colour mean that they are distinctive features, not available in all six TEnTs. This will be further developed in the course of the project.

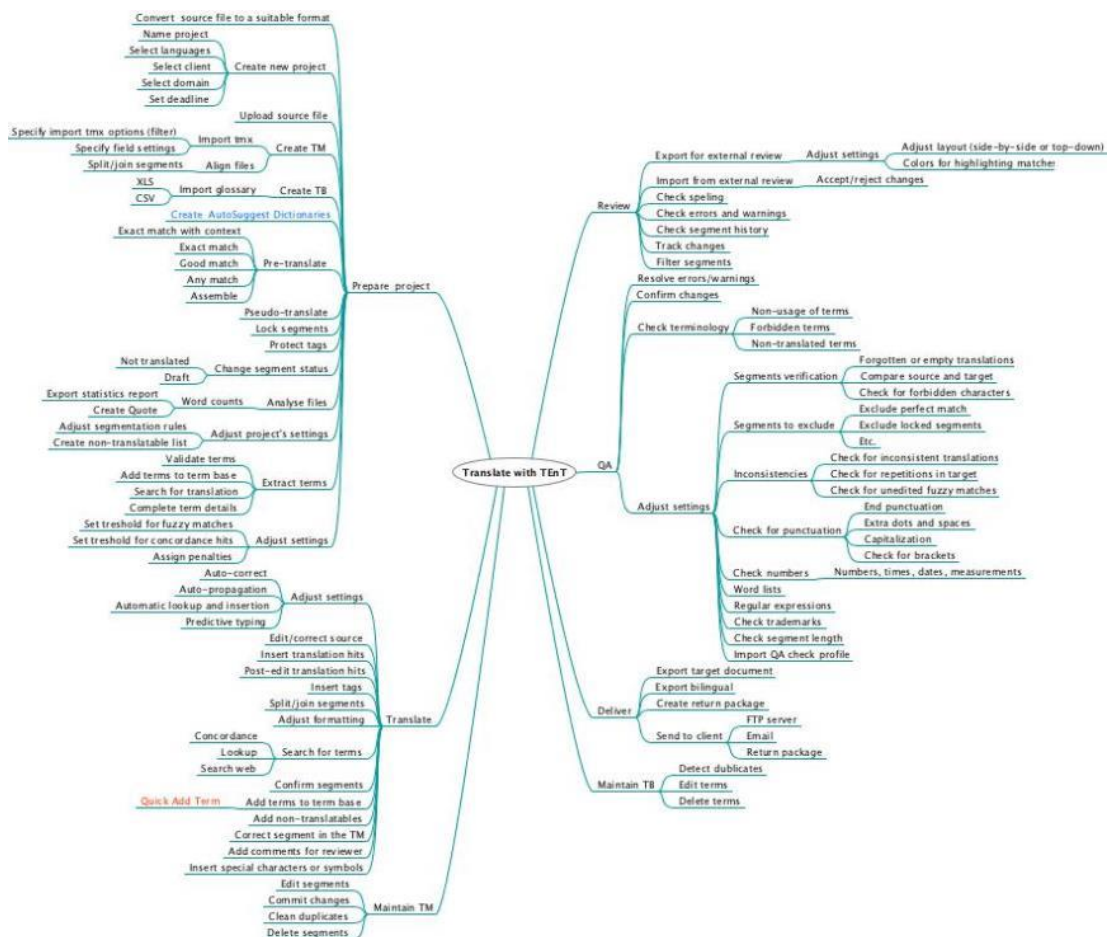


Figure 38 Conceptual map of TEnT features

## Conclusions

We can summarize the preliminary findings of Task 5.1.1 as follows:

- (1) The study of literature revealed that there is a need for more empirical research on ergonomics and usability of translation environments. We believe that the research within WP5 can fill in some gaps in this area.
- (2) The investigation of six representative translation environment tools provides some insights into the TEnT's main modules and features. TEnTs can be either feature-rich or feature-poor, depending on the developer's vision and the target audience.

The traditional client/server-based TEnTs have been designed with the translators' work processes in mind and allow user to follow different workflow scenarios depending on the type project they have: single-file, project workflow, server-workflow, project-packages workflow. For the purpose of this current study, we have looked only at the project workflow during which the user to create a project, upload the files for translation, create new translation memory and import tmx, align bilingual files, create new termbase (or import glossary), translate, and QA.

In the translation editor interface we looked at certain features, such as spell-checker, sub-segment matching, fuzzy match repair, auto-propagation, preview, translation memory, match types, alignment, TM maintenance, terminology features and integration with MT and other resources.

Due to time constraints, we did not look in great details at all features that a TEnT may offer, but the global overview is rendered in the final conceptual map in Figure 38 that may constitute a starting point for further investigations and qualitative studies for teaching and evaluation of translation environments.

The main purpose of the task 5.1.1 was to support the user requirements analysis in Task 5.1.1.2. *Study of work practices of professional translators* that will be presented in the second part of this deliverable. Both the questionnaire and the field research<sup>41</sup> will reveal what TEnTs language professional use and why, what features still need improvement and what types of workflows users may follow during a real-life translation or localization project.

---

<sup>41</sup> By field research we refer both to the interviews and contextual inquiries.

## Study of work practices of professional translators

As the personalization of user interfaces for professional translators is highly related to their work practices, a user-centred approach is encouraged for all tasks related to the UI design and accompanying research. Consequently, we decided to use three techniques to obtain insights in the work practices of professional translators: a widespread web survey to get an understanding of the actual use of current TEnTs<sup>42</sup> by professional translators, semi-structured interviews with companies participating in the IAC of SCATE in order to obtain an overview of workflows in managing translation assignments and contextual inquiries to obtain insights in translators' work practices and workflows. The combination of the three techniques provides a deeper understanding on the use of TEnTs from different perspectives. The following sections present the studies that were conducted and their results.

### *Web survey*

#### Method

In order to get more insight into the experience with translation environment tools (TEnTs) and the general work structure, we designed a questionnaire (Lazar, Feng, & Hochheiser, 2010) aimed at people actively working in the translation environment. The targeted professions were translators (freelance or in-house), interpreters, revisers/editors, post-editors, terminologists and project managers.

The questionnaire contained in total 46 questions related to WP3 and WP5 and were a combination of multiple-choice and open questions. Only the questions related to WP5 will be discussed in this report. These questions are divided into three categories, namely personal information (4 questions), general job information (9 questions) and experience with TEnTs (9 questions).

#### Respondents

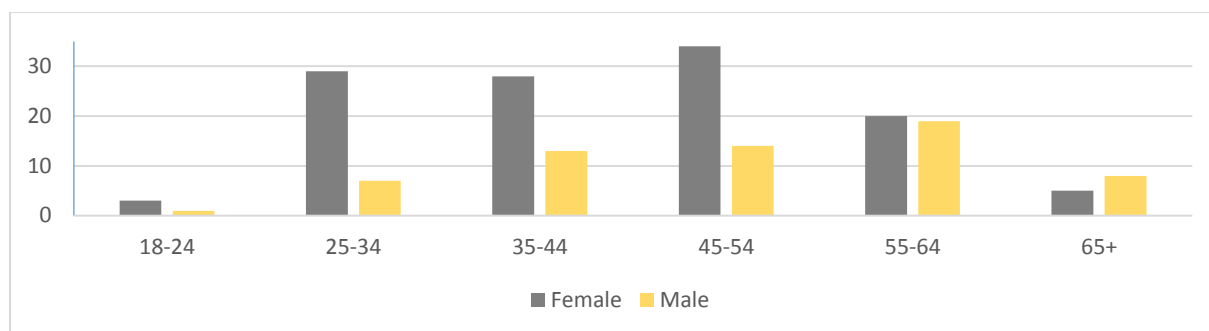
Starting from December 2014 until February 2015, 181 complete and 98 incomplete responses were sent in. The following section describes the results of the complete responses. Incomplete responses were disregarded in the analysis of the results.

---

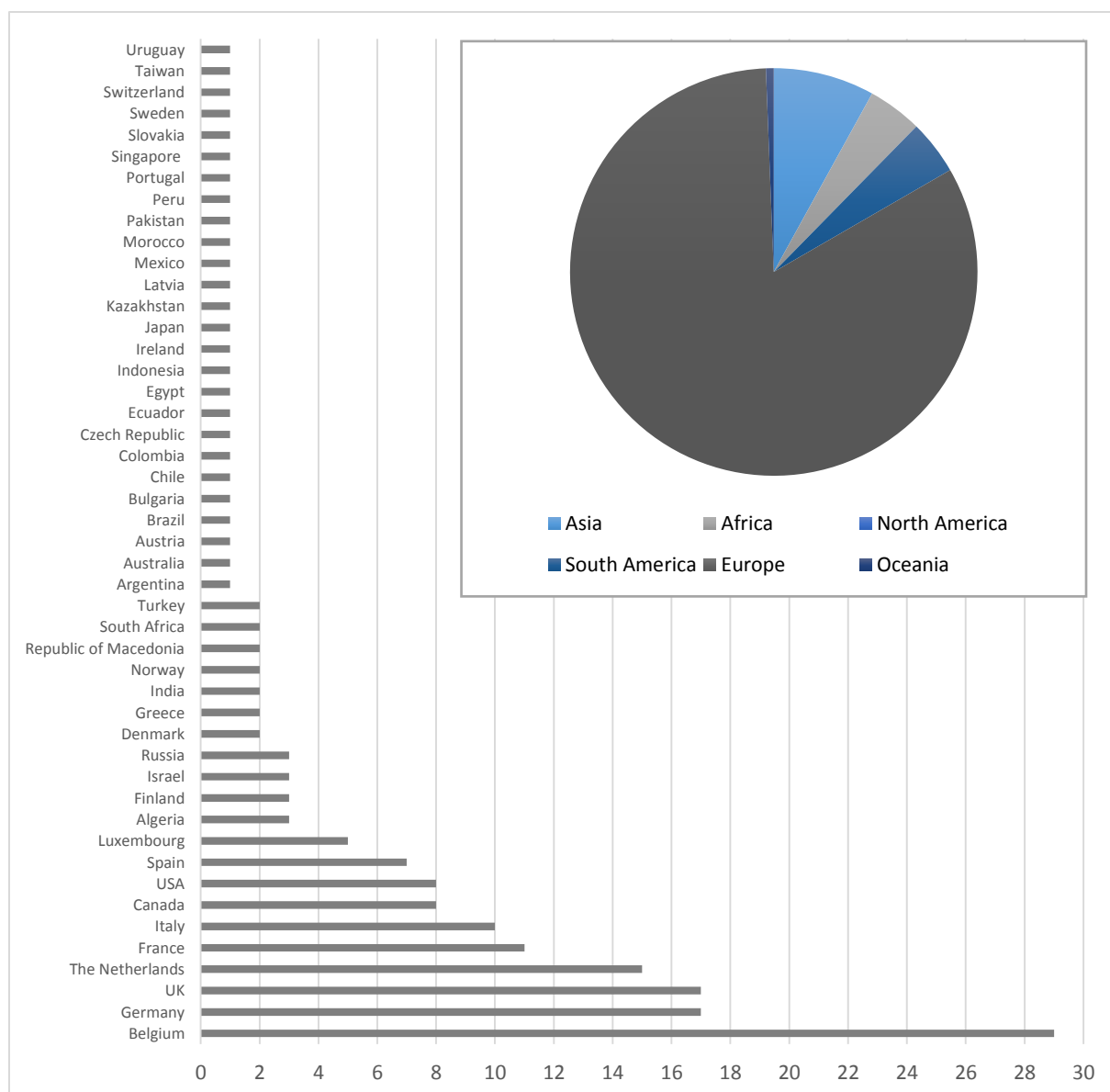
<sup>42</sup> The web survey use the term CAT tools instead of CAT tools. For consistency within this deliverable, we also use TEnT when reporting on the web survey. We believe this is still correct as CAT tools are TEnTs.

### Demographics

A large proportion of the 119 female and 62 male respondents ranged in age between 25 and 64, with 4 respondents aged between 18 and 24 and 13 respondents aged 65 or older (see *Figure 39*). Figure 2 shows that respondents originated from a total of 47 different countries. Europe was largely represented (74,86%), followed by North-America (9,5%), Asia (7,26%), South-America (3,91%), Africa (3,91%). Only one respondent indicated Australia (Oceania) as their country of residence. The most frequently occurring countries were Belgium (16,20%), followed by Germany (9,50%), The United Kingdom (9,50%), The Netherlands (8,38%), France (6,15%), Italy (5,59%) and the United States of America (4,47%).



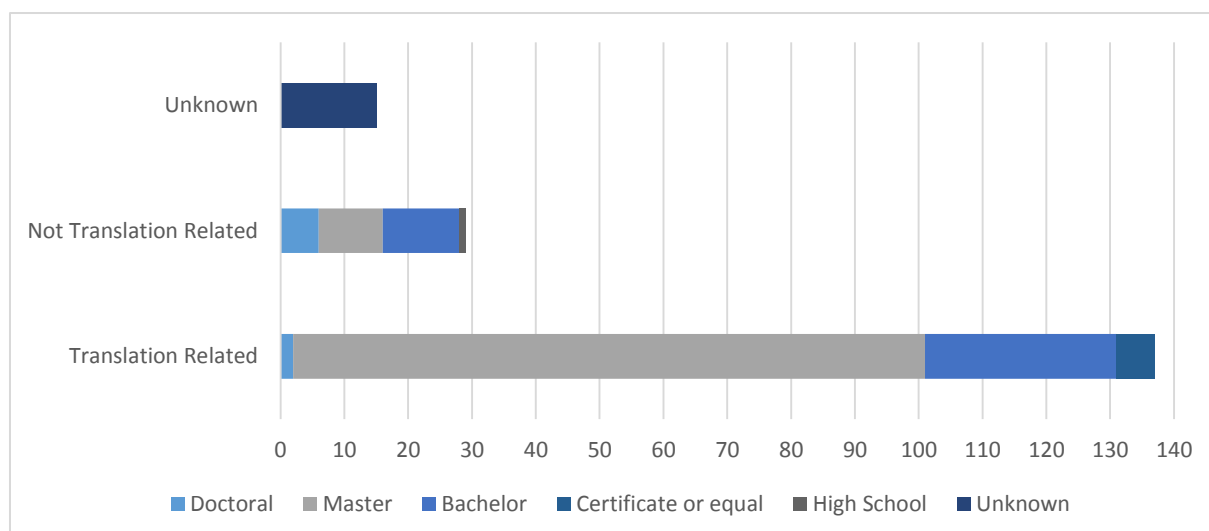
*Figure 39 Gender distribution per age group*



*Figure 40 Country distribution (Inset: Continental distribution)*

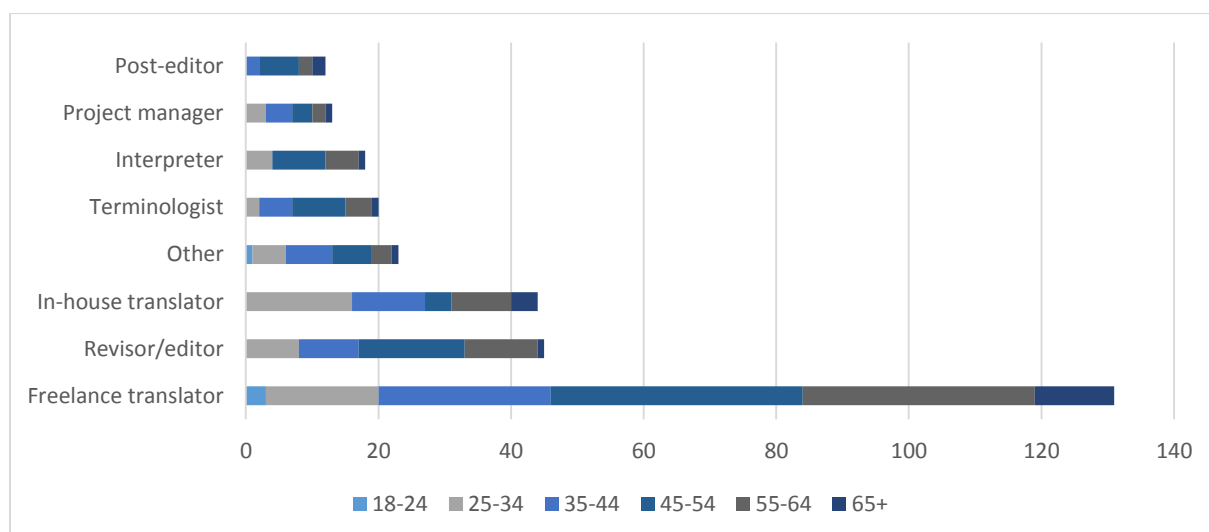
Respondents with a degree in translation (75,69%) were predominant. The most common level of degree was a master's degree (54,70% translation related, 5,52% not related to translation, 60,22% total), followed by a bachelor's degree (16,57% translation related, 6,63% not related to translation, 23,20% total) and a PhD (1,10% translation related, 3,31% not related to translation, 4,42% total). A certificate in translation was indicated by 6 respondents (3,31%) and one respondent indicated a degree of high school (0,55%). 15 respondents (8,29%) did not indicate or clarify their level of degree.





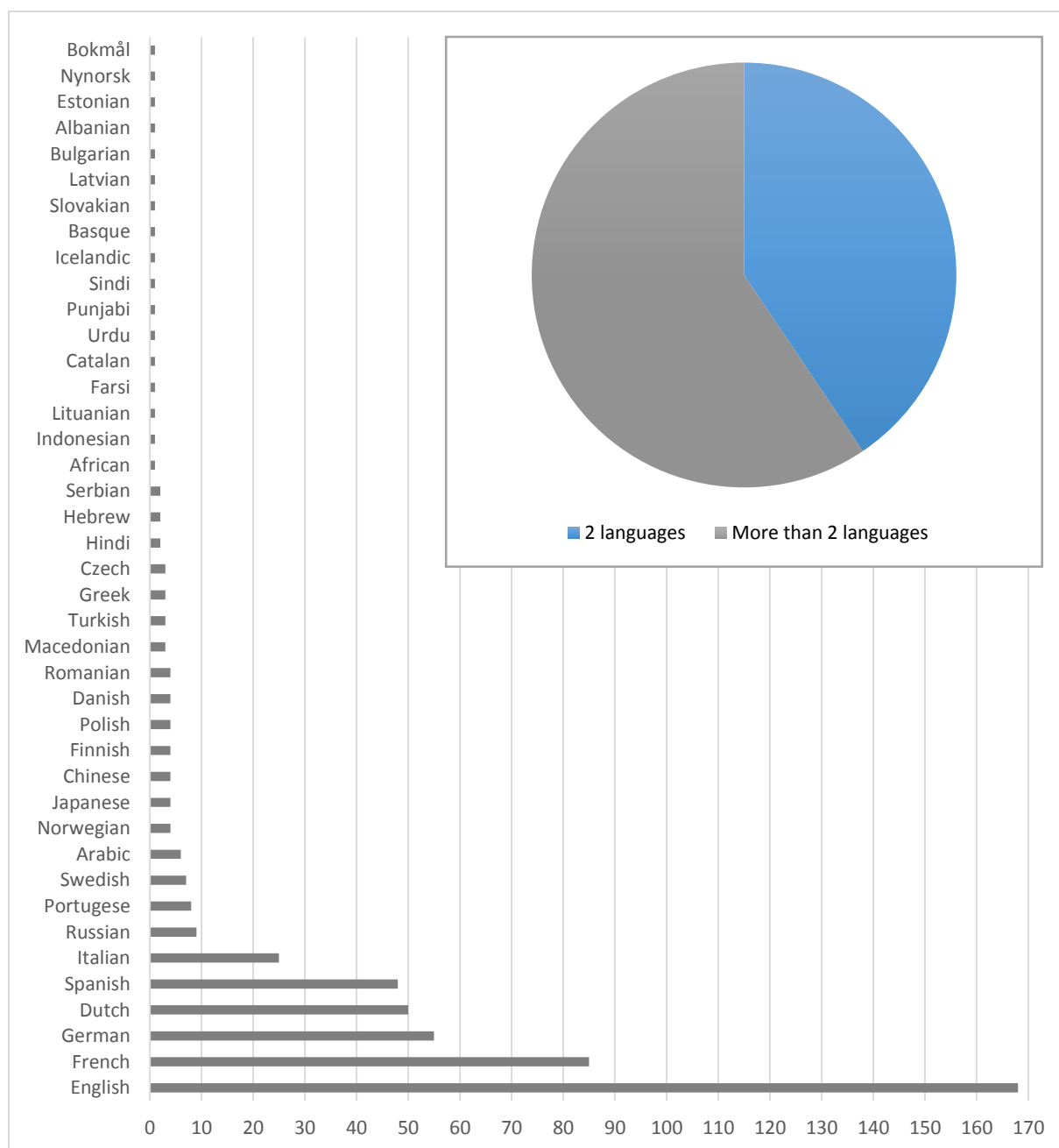
*Figure 41 Degree levels*

The most common profession was that of freelance translator (72,38%), followed by reviser/editor (24,86%) and in-house translator (24,31%). Less common job descriptions for participants were terminologist (11,05%), interpreter (9,94%), project manager (7,18%) or post-editor (6,63%). Respondents also indicated other tasks (12,15%) as their job description, such as researcher, professor, trainer, student, engineer and lecturer.



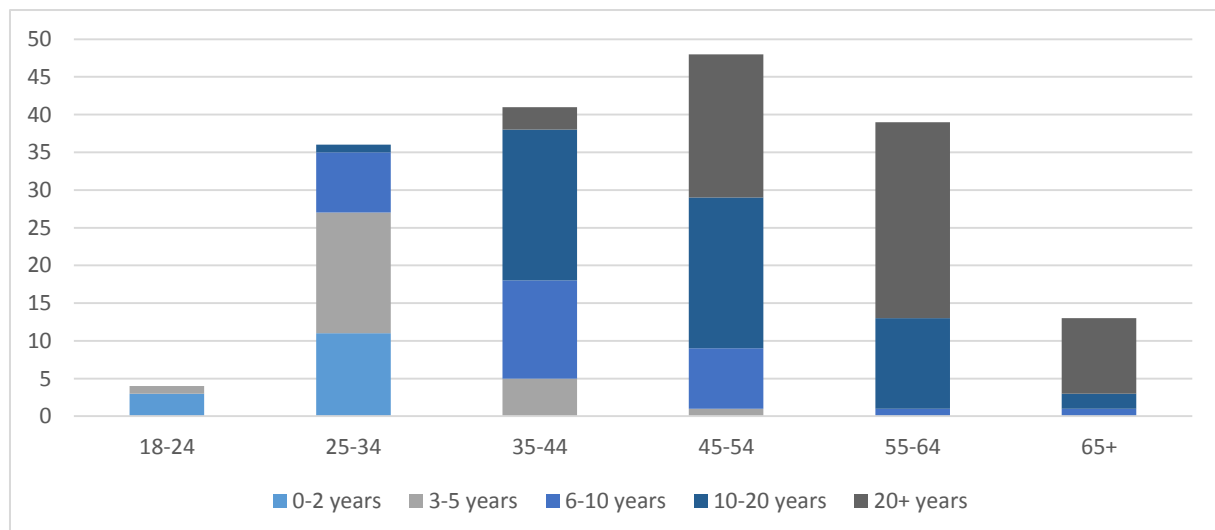
*Figure 42 Age group per job description*

Respondents were asked which language pair(s) they worked with. A total of 41 different languages were registered. The most common languages were English (32,06%), French (16,22%), German (10,50%), Dutch (9,54%) and Spanish (9,16%). Most respondents work with more than two languages (59,12%), compared to working with two languages (40,33%). One respondent did not indicate a language pair (0,55%).



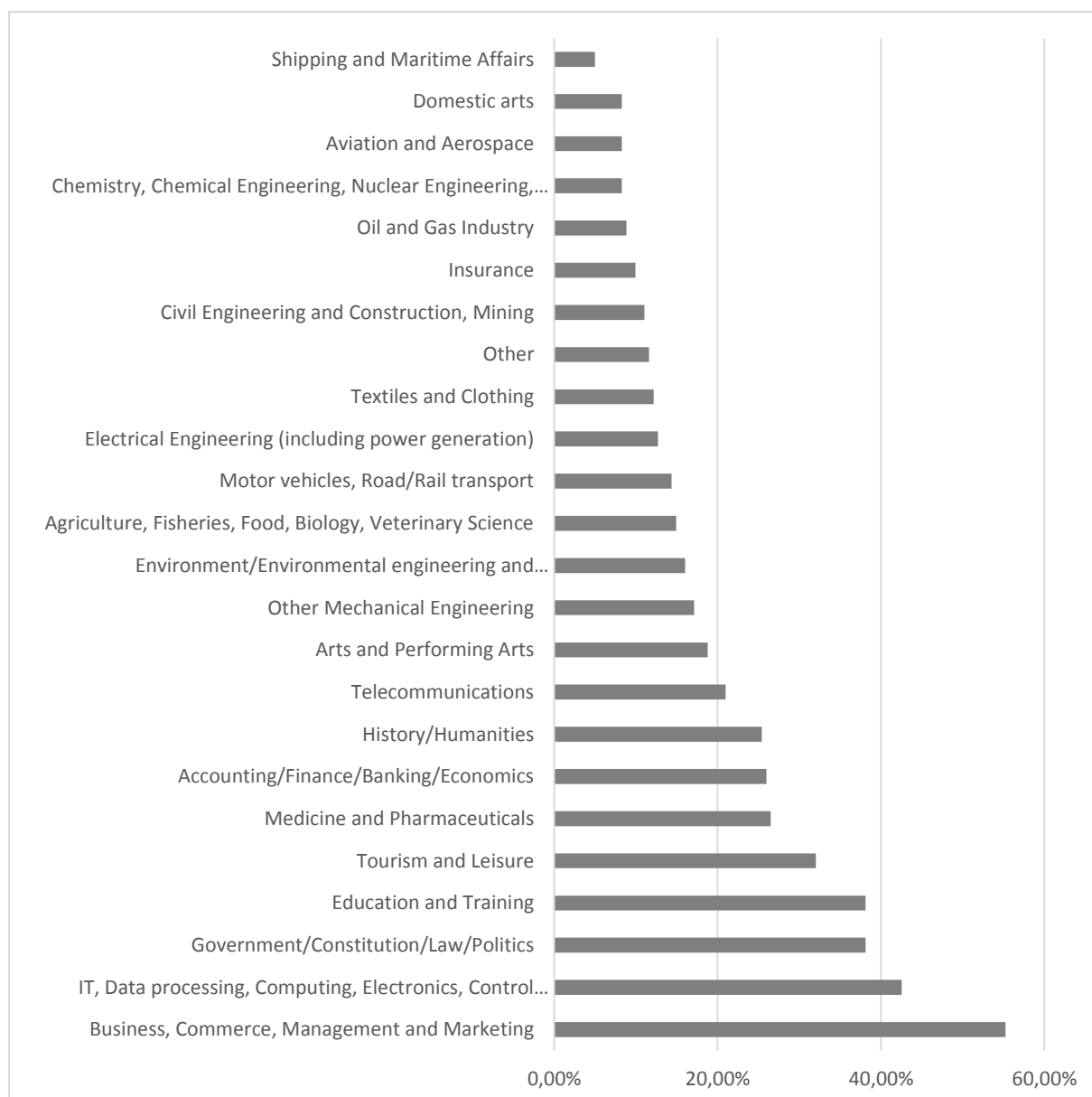
*Figure 43 Language frequency in language pairs (Inset: Respondents working with 2 languages compared to working with more than 2 languages)*

A considerable amount of respondents indicated a work experience of 10 years or more (30,39% 10–20 years, 32,04% more than 20 years, 62,45% total). The 14 participants (7,78%) that indicated to have less than 2 years of work experience, can be found in the lower age groups 18–24 and 25–35. Older respondents showed an increasing amount of experience.



*Figure 44 Work experience per age group*

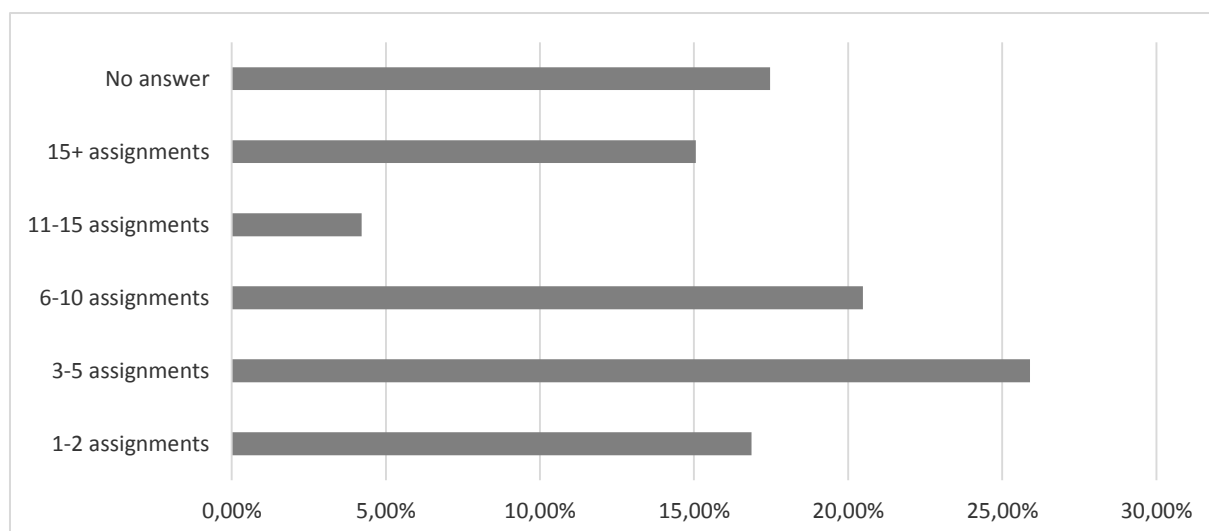
Respondents were asked to indicate one or more fields of expertise (see *Figure 45*). On average, respondents picked 4,83 fields of expertise with 21 respondents picking only one and 1 respondent picking 22 fields. More than half of the respondents indicated a field of expertise related to Business, Commerce, Management and Marketing (55,25%), followed by IT, Data processing, Computing, Electronics, Control Engineering (42,54%). The shared third place was for Government/Constitution/Law/Politics (38,12%) and Education and Training (38,12%). The least picked field was Shipping and Maritime Affairs (4,87%). Respondents also indicated not listed fields (11,60%), such as Literature, Film Scripts, Mining/Geology and Amusement Parks. Two respondents indicated no field of expertise as they just started working or were still studying.



*Figure 45 Fields of expertise*

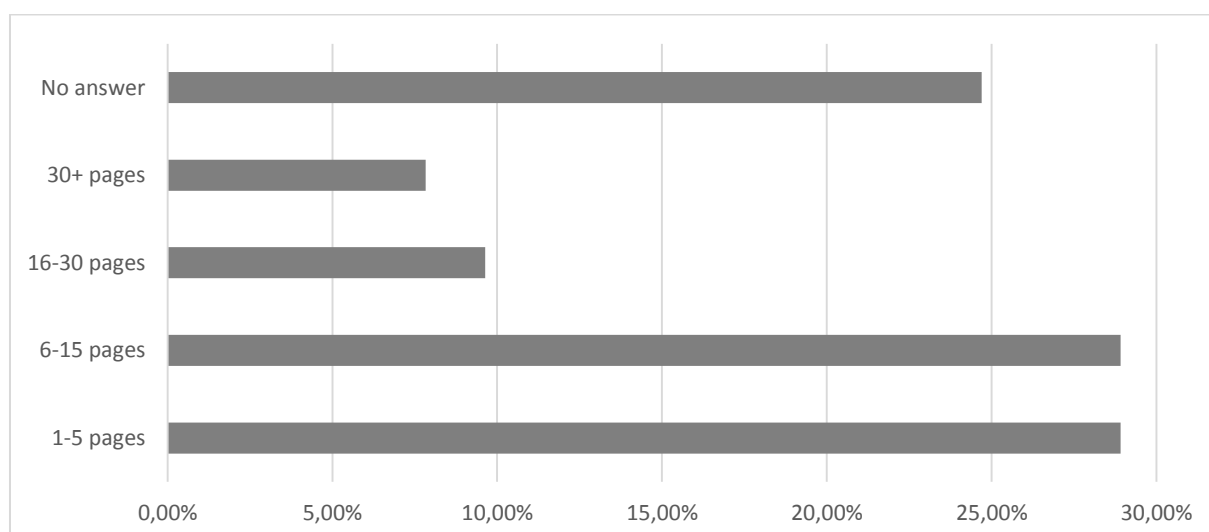
Three conditional questions were only presented to respondents if they indicated to have a profession of either translator (freelance or in-house), post-editor or terminologist. These questions inquired about the amount of assignments they complete on average each week, the average number of pages per assignment and the average amount of words they translate each day.

Figure 46 shows that a large amount of respondents indicated an average of 1–2 translation jobs per week (16,87%), 3–5 translation jobs per week (25,90%) or 6–10 translation jobs per week (20,48%). Only 7 respondents indicated an average of 11–15 translation jobs per week (4,22%) and 25 indicated an average of 15 or more translation jobs per week (15,06%). 29 respondents (17,47%) did not answer the question.



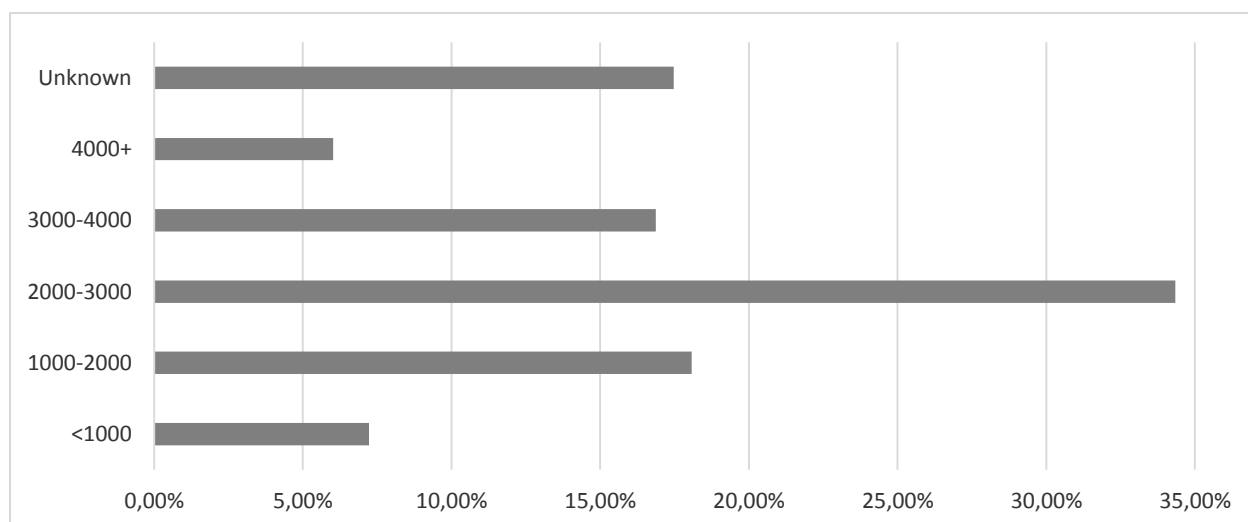
*Figure 46 Average amount of translation assignments per week*

For most respondents, the average number of pages per assignment ranges between 1 and 15 pages (1–5: 28,92%; 6–15: 18,92%). A large amount of respondents (24,70%) did not indicate an average (Figure 47).



*Figure 47 Average amount of pages per assignment*

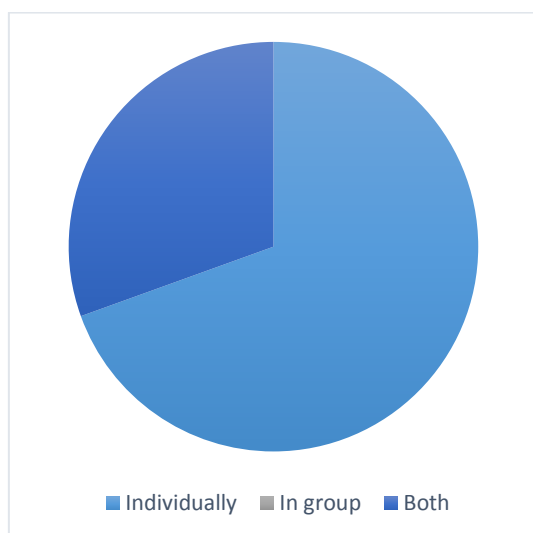
Figure 48 shows that on average, a large amount of respondents translates between 2000 and 3000 words each day (34,34%), followed by 1000 to 2000 words (18,07%) and 3000 to 4000 words (16,86%). 12 respondents indicated to translate less than 1000 words each day (7,23%) and 10 indicated more than 4000 words (6,00%). 29 respondents (17,47%) did not provide an average. Comments made by the respondents included that the average amount of words “*depends on the type of material that is to be translated and the urgency*” or “*varies too much*”. One respondent commented that “*word-count is an irrelevant indicator when comparing Fenno-Ugric and Indo-European languages*”.



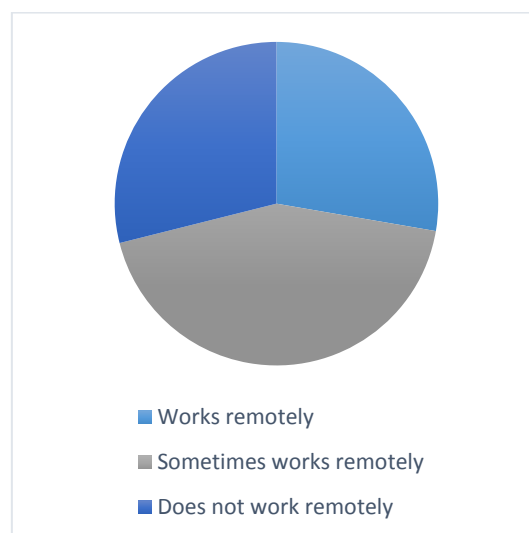
*Figure 48 Average amount of translated words each day*

Respondents were asked if they usually translate assignments individually or in group (see Figure 49). This question was only presented to respondents with the profession of in-house translator, post-editor or terminologist. 69,49% of these respondents mostly work individually while 30,51% of the respondents work both individually or in group. None of the respondents indicated to work solely in a group context.

Respondents with the profession of in-house translator, reviser/editor, post-editor, terminologist or project manager were asked if they work remotely. Most respondents indicated that they sometimes work remotely (43,97%) next to not working remotely (28,92%) and working remotely all the time (27,71%).



*Figure 49 Distribution of group work*



*Figure 50 Distribution of remote work*

### Experience with TEnTs

The usage of translation environment tools is prevalent (76,80%). This indicates the added value of such tools in the work environment is clear. The majority of the respondents that indicated not to use TEnTs, can be situated in the age groups 55–64 and 65+ (see Figure 51).

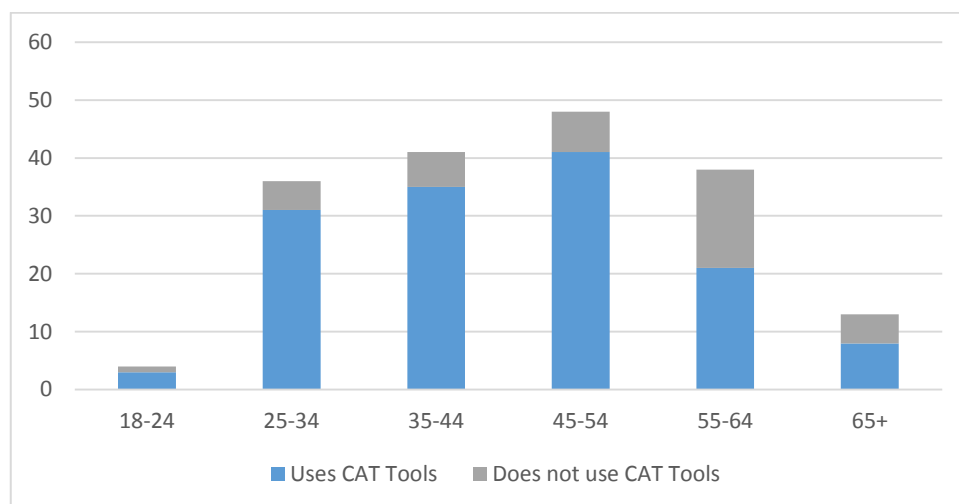


Figure 51 Usage of TEnTs per age group, lower parts of the bars indicate TEnT usage.

Most respondents that use TEnTs, have an experience of 10 years or more (38,13%), followed by 6–10 years (23,02%), 3–5 years (22,30%) and 0–2 years (16,55%). Figure 52 shows the usage of individual tools. Respondents could choose multiple tools. The most used tool is SDL Trados Studio (71,22%), followed by memoQ (26,62%). These tools are mostly used full-time (69,78%). Part-time usage was indicated by 31 respondents (22,30%) and occasional usage by 11 (7,91%).

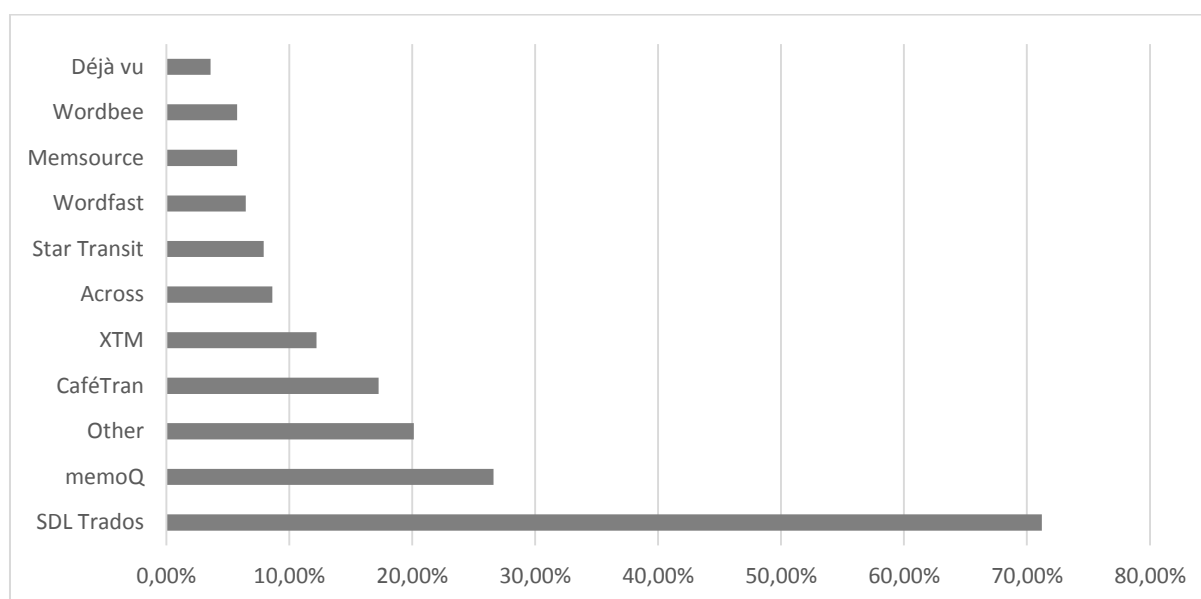
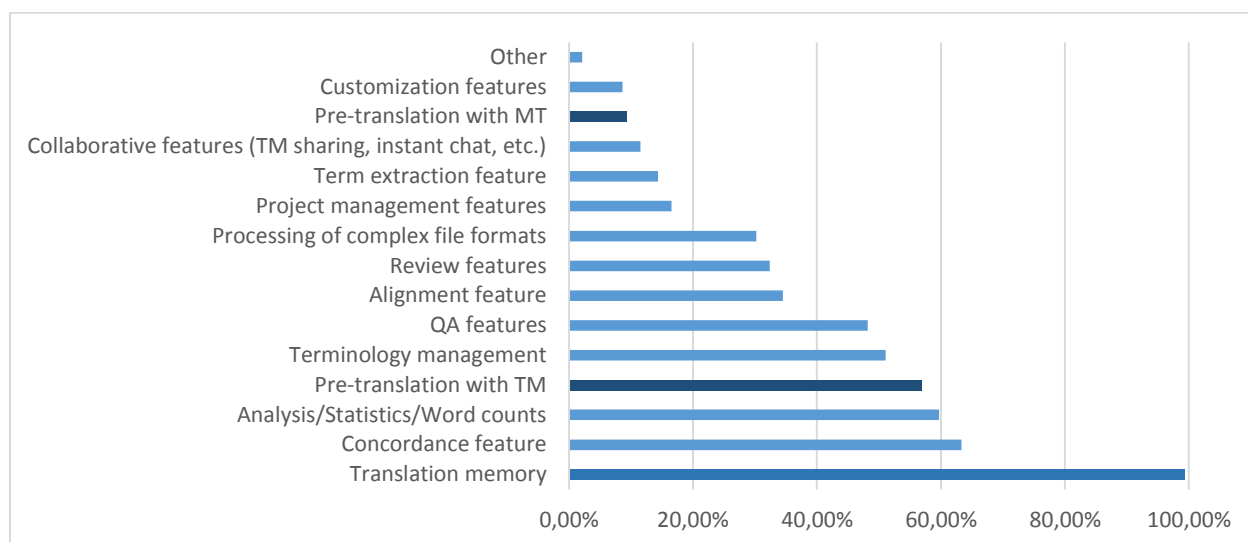


Figure 52 Used TEnTs

Nearly all respondents that used TEnTs, used the translation memory (see *Figure 53*). Other widely used features are concordance (63,31%), analysis/statistics/word counts (59,71%), pre-translation with TM (56,83%), terminology management (51,08%) and QA features (48,20%). Less used features include alignment features (34,53%), review features (32,37%) and processing of complex file formats (30,22%). The least used features are project management features (16,55%), term extraction features (14,39%), collaborative features (TM sharing, instant chat, etc.) (11,51%), pre-translation with MT (9,35%), customization features (8,63%) and others (2,16%).



*Figure 53 Most frequently used features of TEnTs, translation memory is used nearly always.*

Figure 54 shows the most used features per job category. Usage of the feature is however not always related to the job category itself as respondents could indicate multiple job categories. Other features are used much less by specific job categories: statistics (in-house translators), concordance (interpreters).



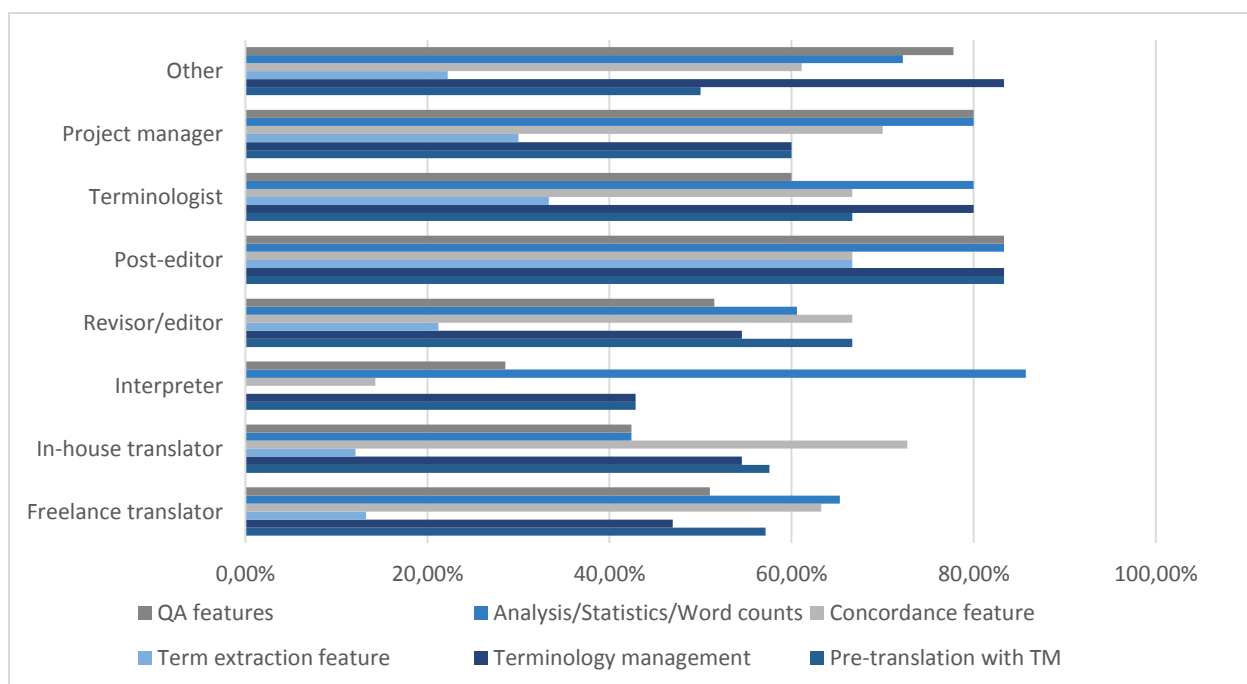


Figure 54 Most frequently used features per job description relative to usage of translation memory

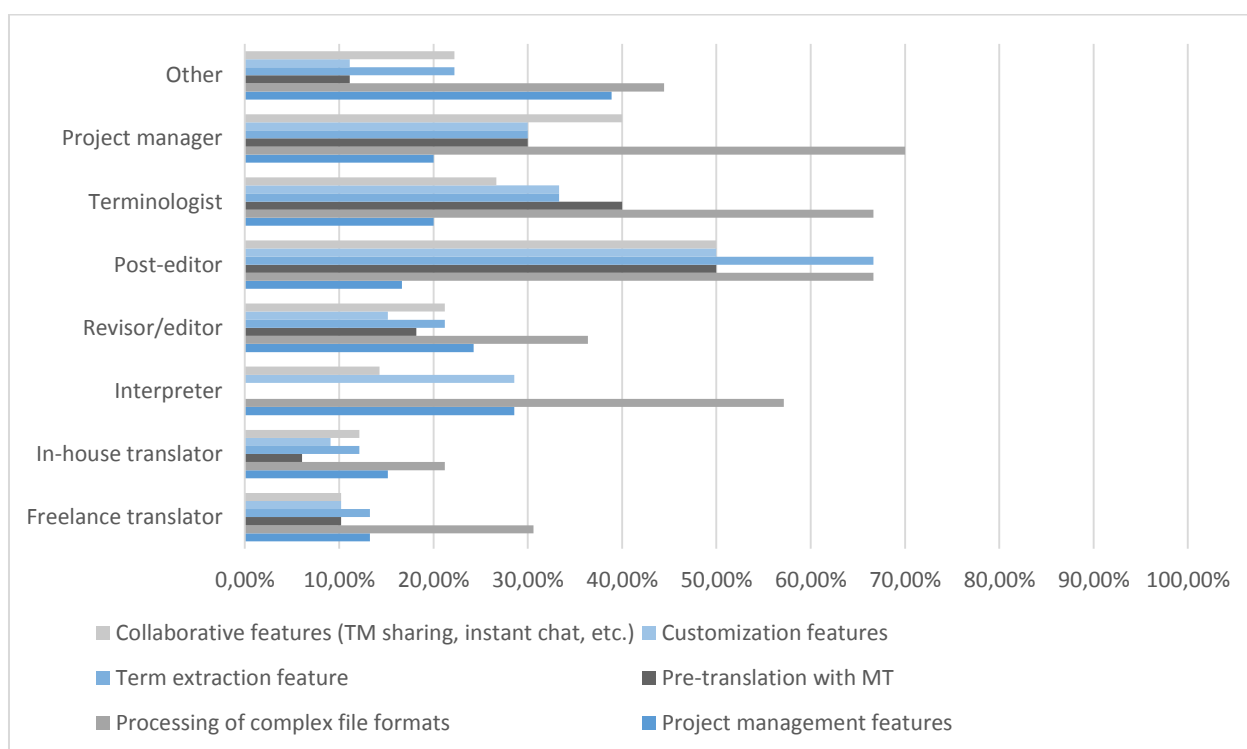


Figure 55 Least used features per job category

Figure 55 shows the least used features per job category. This reveals that some features are used less because they are not used equally much by all job categories, and especially the free-lance and in-house translators, which are the dominant group. Term extraction is used relatively much by people also doing post-editing (2/3 of 12 respondents in this job

category use this feature) while for other job categories this is at most 40%. Processing of complex file formats is done much less by translators and revisers/editors (< 40%), while about 70% of the project managers, post-editors and terminologists, and almost 60% of the interpreters use this feature. Project management is lowly used across all named job categories (<30%), but this might be caused by the nature of the support provided as 60% of XTM users use this feature (see Figure 57). The fact that all other lowly rated features are used by at least 50% of the post-editors seems interesting, but may be partially due to the low number of people with this job category (12) and the fact that a single respondent could give multiple job categories.

When looking at project management features for each TEnT, 60% of XTM users use these features compared to 14,43% of SDL Trados Studio users.

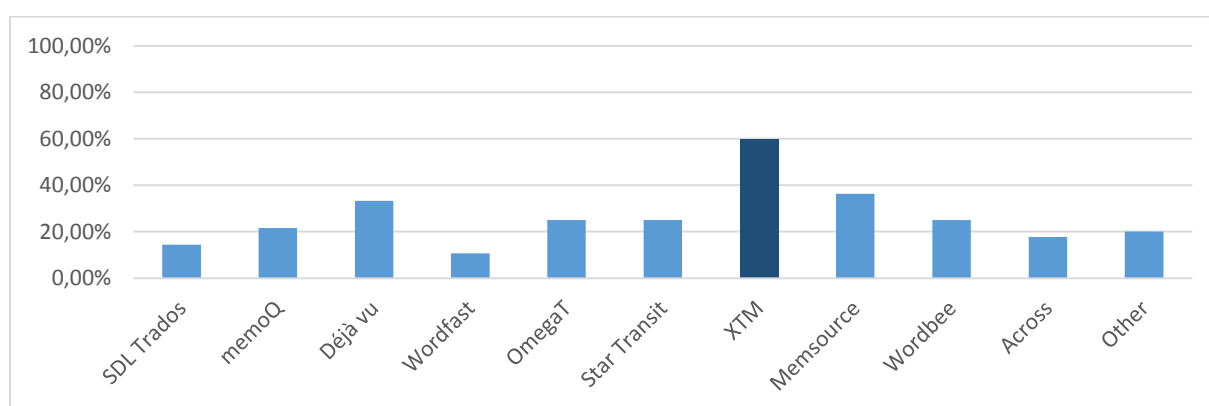


Figure 56 Project management features per TEnT

When comparing pre-translation features with MT and TM across TEnTs, users of Star Transit score the highest (87,50% TM, 25,00% MT) as can be seen in Figure 57.

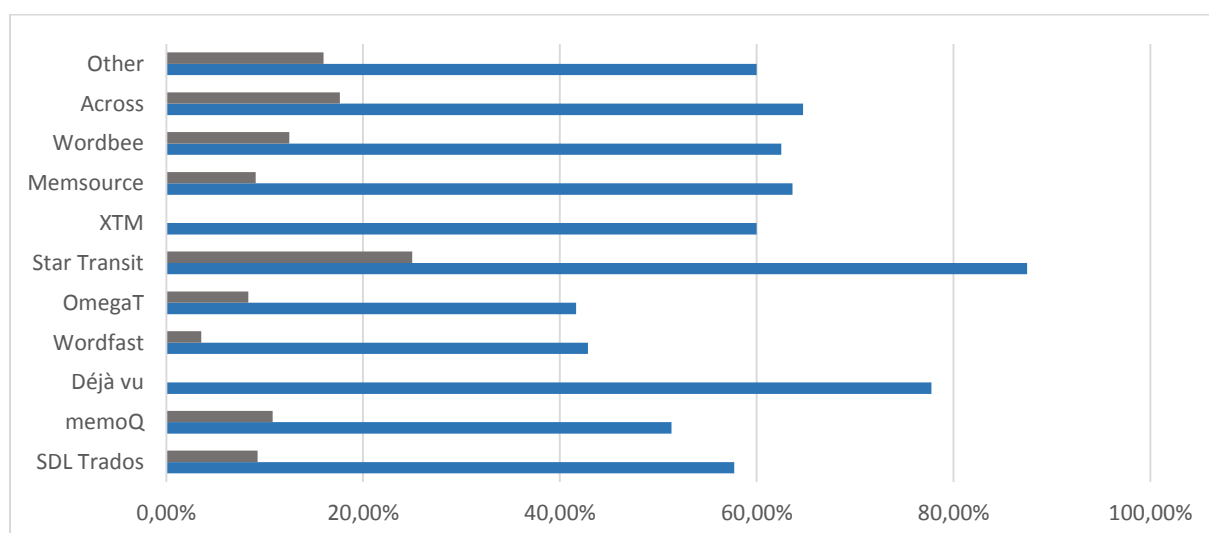
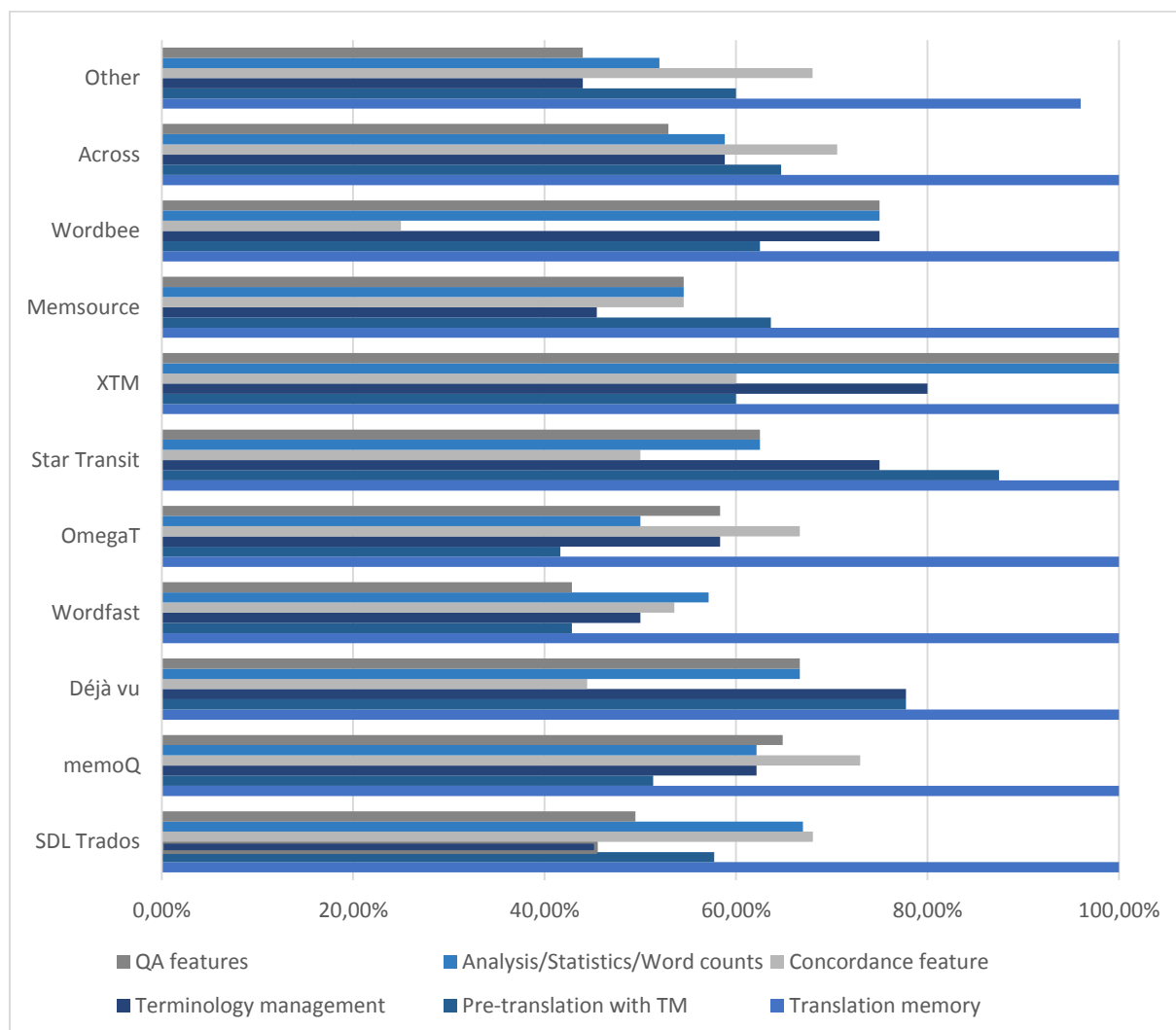


Figure 57 Pre-translation with MT and TM per TEnT

Figure 58 shows the most used features across TEnTs. XTM users also heavily rely on QA features and analysis/statistics/word count. Terminology management in SDL Trados Studio users is the lowest compared to other TEnTs.



*Figure 58 Most frequently used features per TEnT*

Figure 59 gives an overview of the tool usage split out per job category. The graph shows that SDL Trados dominates all categories except interpreters, where there is a tie with Wordfast a bit above 20%. Wordbee and XTM also have very specific audiences, respectively mostly translators and project managers.

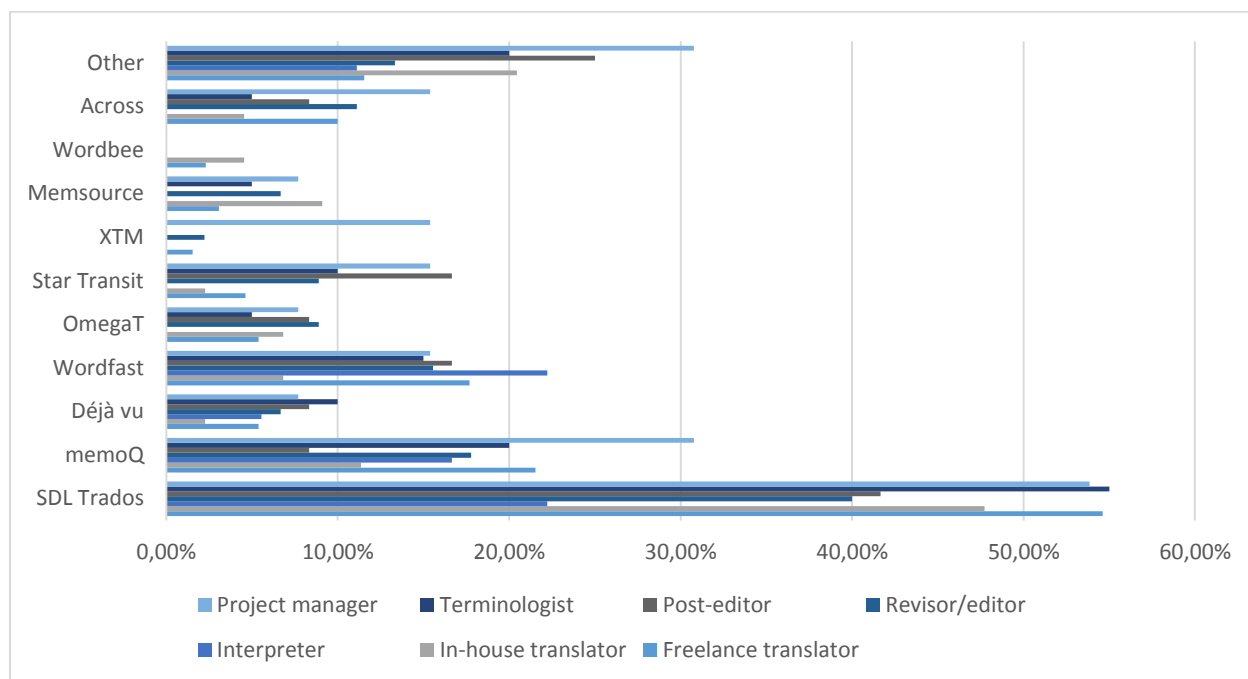


Figure 59 TEnTs per Job category, the number of users of freelance translators

Users of TEnTs prefer these tools because of terminology consistency (92,09%), time efficiency (83,45%), productivity increase (79,86%), quality improvement (60,43%) and ease of collaboration (30,22%). Other answers indicated a preference because of layout preservation (1 respondent), QA tools (2 respondent), easier to work in one single environment (1 respondent) and customizability with regex (1 respondent). One respondent commented that the use of TEnTs is a “*work prerequisite*” and another commented he does not prefer working with TEnTs (“*I don’t prefer it!*”).

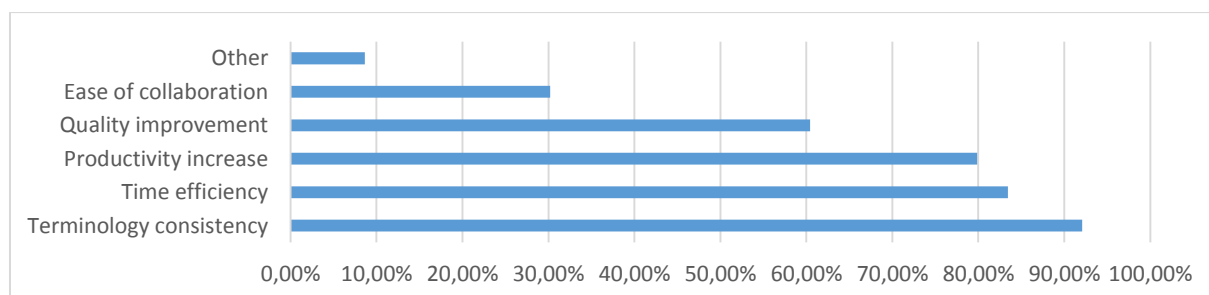
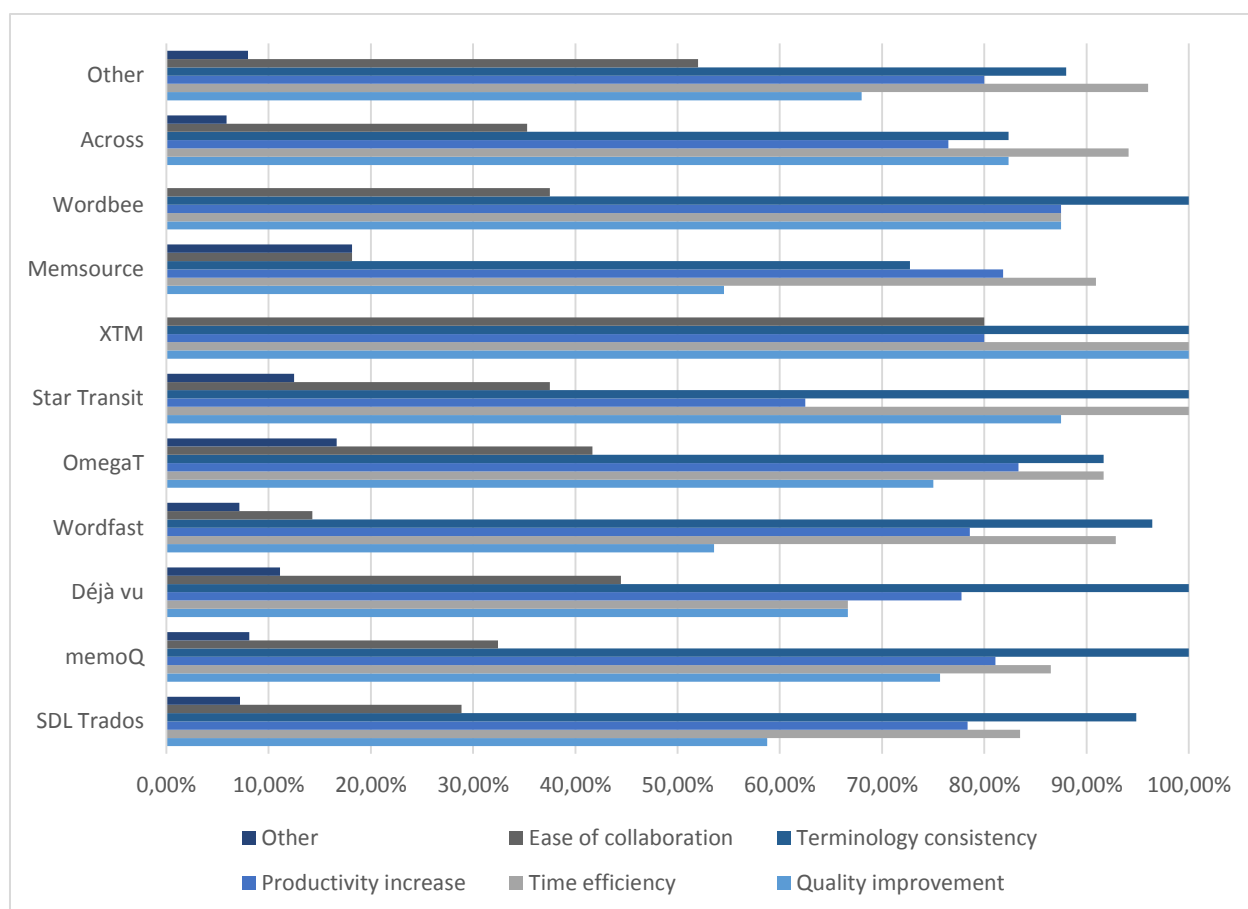


Figure 60 Reasons respondents prefer to work with TEnTs

When comparing these reasons across TEnTs, terminology consistency is clearly indicated by users of all TEnTs. Besides terminology consistency, XTM is also preferred for time

efficiency and quality improvement. Ease of collaboration is rated the lowest by users of SDL Trados Studio.



*Figure 61 Reason for preference of a TEnT per TEnT*

Respondents that use TEnTs were asked which features needed to be improved or added in order to optimize their translation process. 73 respondents answered this open question, which resulted in 24 features. The most important feature mentioned it the support for terminology management (10,79%), followed by translation features/language support/MT (8,63%), tools & integration (7,19%), file format support (6,47%), user-friendly UI/WYSIWYG UI (6,47%), project management (4,32%) and customization features (3,60%).



*Figure 62 TEnT features that need improvement. Features that may be related to WP5 are highlighted*

When comparing TEnT features that need improvement across different TEnTs, the users of Déjà Vu are the highest to request improvement in terminology management. Star Transit and Across users would like improvements in file format support. Wordbee, XTM and Memsource users request improvements regarding project management features. Requests to improve translation features, language support and machine translation are found across all tools.

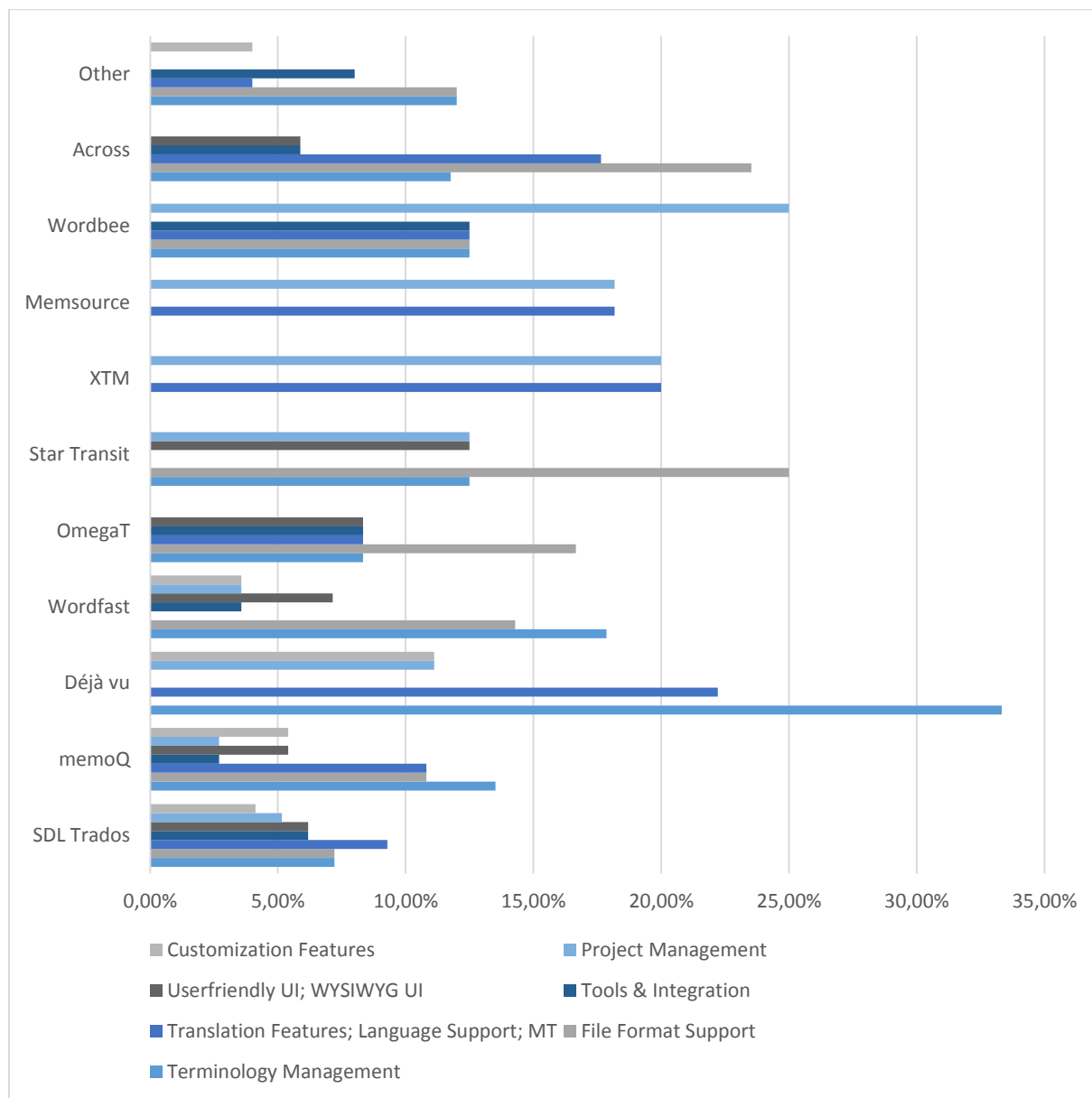


Figure 63 Top TEnT features that need improvement per TEnT

A considerable amount of users of TEnTs indicated that they learned these tools by themselves (78,42%), followed by (in-house) training (34,53%), video tutorials (33,09%) and webinars (30,22%). Only few respondents received TEnT training during their education (17,27%), workshops (15,83%) or other means of learning (5,04%).

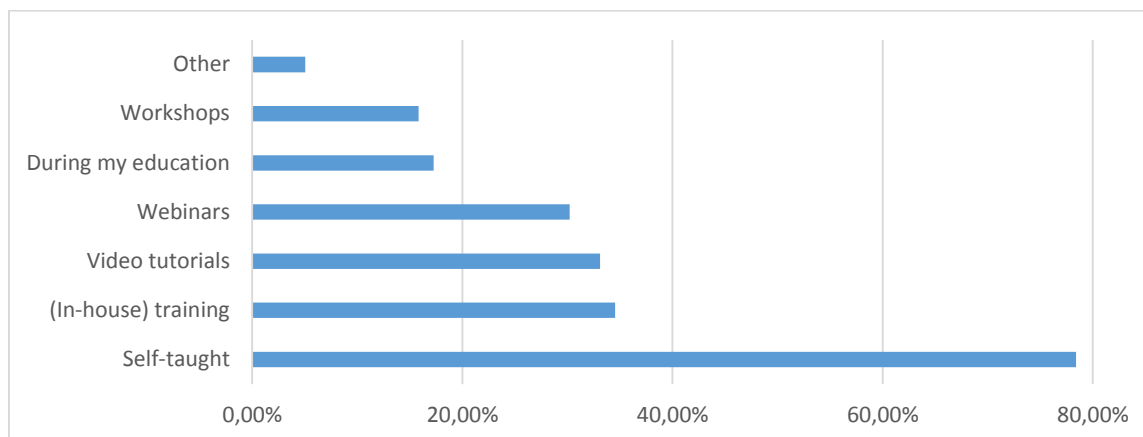


Figure 64 How respondents learned to use TEnTs

When comparing how respondents learned to use TEnTs per TEnT (see Figure 65), users heavily rely on self-teaching. Users of Star Transit use video tutorials more than other users, but severely lacks workshops. XTM and Wordbee are taught more during the user's education than other users. Memsources, XTM, Wordbee and Star Transit users receive the largest amount of in-house training.

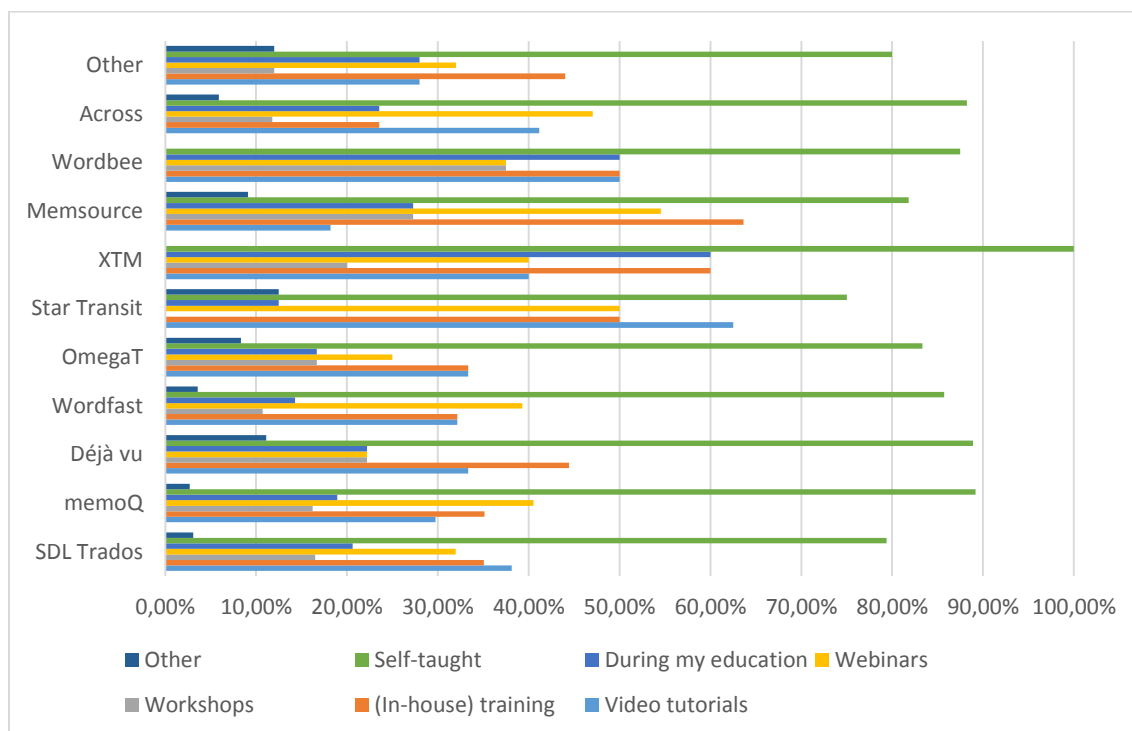


Figure 65 How respondents learned to use TEnTs per TEnT



Respondents that use TEnTs could indicate multiple aspects they found important when using a TEnT (see). 86,33% of the respondents marked the importance of a tool being easy to use, followed by resource management (69,78%), speed (65,47%), ease of learning (64,75%), compatibility with other tools (58,27%), quality assurance checks (51,80%) and easy to customize (39,57%). Project management (7,91%) is considered a less important aspect. Respondents also had the option to add other aspects, which included “*support requirements*”, “*ergonomic use*” and “*stability of the tool*”.

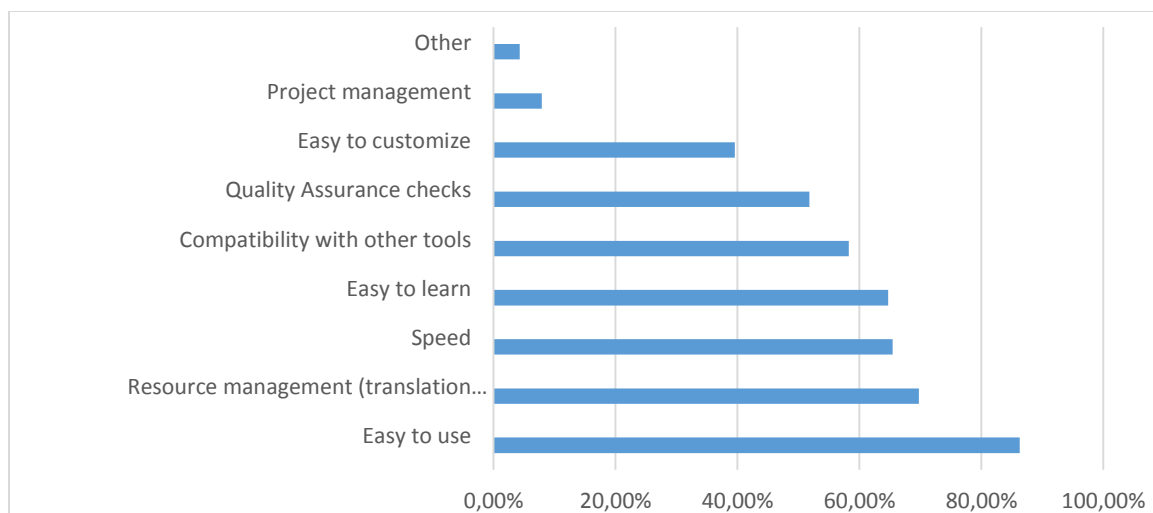


Figure 66 Important aspects when using TEnTs

When comparing important aspects to the TEnTs respondents use, XTM users rate speed (100,00%) the highest. Ease of use, compatibility with other tools, resource management, QA checks and ease of learning are considered valuable aspects regardless of the TEnT used. Project management is generally rated lower, but highest by users of XTM, Memsource and Déjà Vu. Ease of customization is an important aspect for all TEnT users, but rated lowest by users of Memsource.

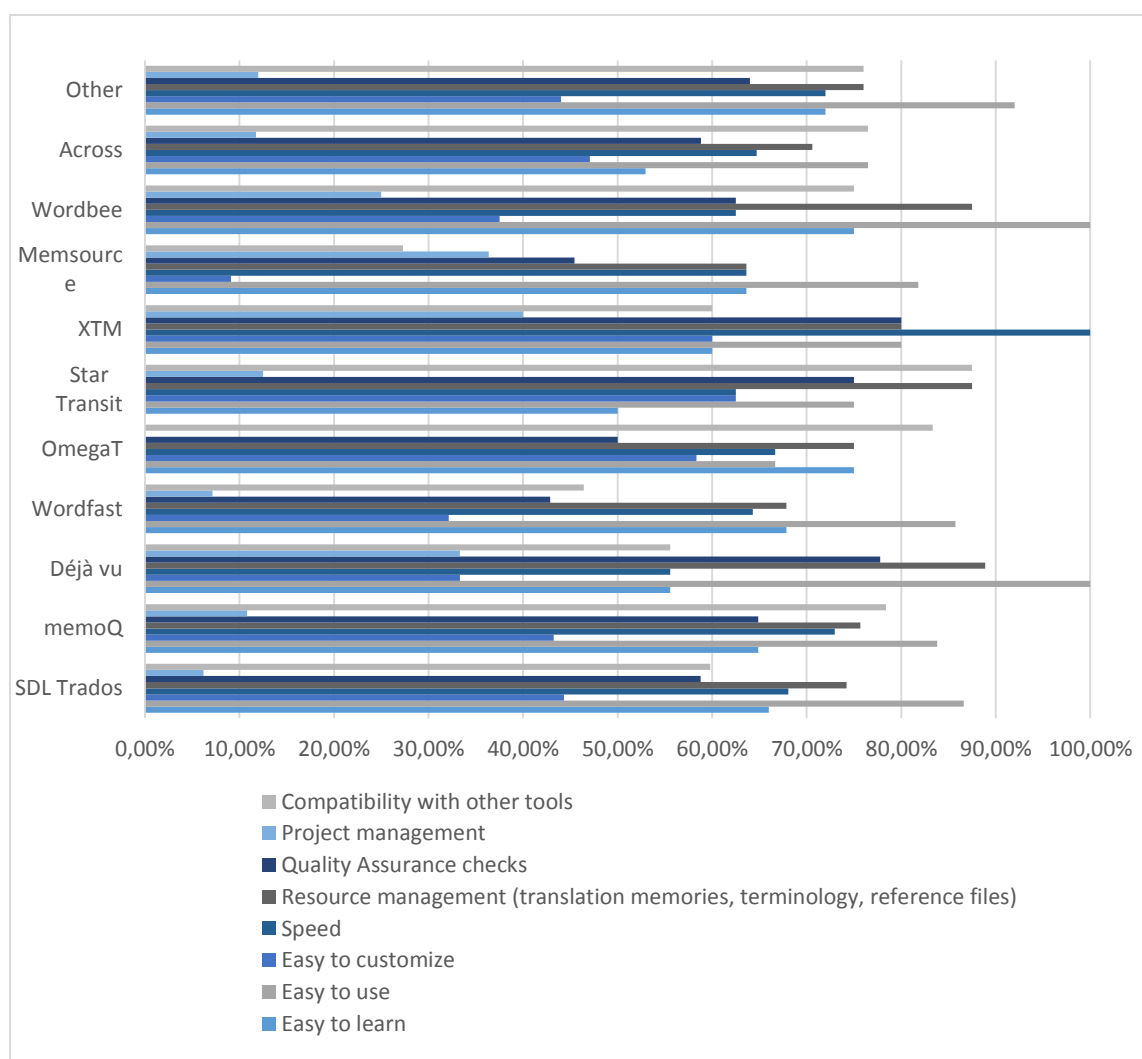


Figure 67 Important aspects when using TEnTs per TEnT

## Conclusions

In order to get more insight into the experience with translation environment (TEnT) tools and the general work structure, we designed a questionnaire aimed at people actively working in the translation environment. A total of 22 questions related to personal information (4 questions), general job information (9 questions) and experience with TEnTs (9 questions) received 181 completed responses.

The most important findings of the questionnaire can be summarized as follows:

Language professionals use TEnTs because of terminology consistency, time efficiency, productivity increase, and quality improvement. The criteria for selecting a TEnT may vary per user profile and needs. However, important criteria for selection include ease of use, resource management, speed, ease of learning, compatibility with other tools, QA checks and ease of customization. The project management features and ease of collaboration are more important criteria for the project managers.

As far as the TEnT training is concerned, users mainly learn how to work with a tool by themselves. In comparison with the freelance translators, the in-house translators may benefit from in-house trainings and workshops.

With regard to TEnT usage, SDL Trados Studio is still a market leader, closely followed by memoQ, CafeTran and XTM. Most used features are translation memory, concordance, statistics and word count, pre-translation with TM, QA and terminology management. Some of the features may be used in one TEnT more than in others depending on their level of implementation. For example, the Déjà Vu users may use the pre-translate feature with MT more than other TEnT users due to the unique feature of Déjà Vu, namely the TM fuzzy match repair feature.

These results provide insights in the specifics and preferences of a wide range of TEnT users. In order to obtain a better understanding of why language professionals have particular preferences, we need qualitative studies to complement this web survey. The following sections present the results the qualitative studies conducted in the SCATE project, semi-structured interviews and contextual inquiries, which provide insights into the work practices and workflows of different user-profiles.

## ***Semi-structured Interviews***

Besides the web survey, we conducted semi-structured interviews (Lazar, Feng, & Hochheiser, 2010) with companies that are involved in translation on a daily basis. Semi-structured interviews are used to understand the user needs based on a series of interviews. In a semi-structured interview, the discussion starts with a set of fixed questions but allows to freely discuss topics that come up during the interview. In order to prepare the fixed set of questions, we asked the companies to fill in a short questionnaire. This questionnaire provided us with insights in the companies' particular interests in SCATE, their daily business and the possibility to be involved in our studies, and inspired us for further questions that could be asked during the semi-structured interviews. The semi-structured interviews took place between October 2014 and January 2015. We visited five companies. During the interviews, the general workflow of each organization, the type of translators they work with and the translation software they use, were discussed. These interviews resulted in a workflow that represents how the organisations involved in our interviews handle translation assignments. Because of confidentiality reasons, we do not include the reports of these interviews but we immediately present the resulting workflow.

## **Workflow**

We extracted a generalized *workflow* from the aforementioned interviews with the companies. This workflow contains the basic steps to fulfil a translation request including all the different *actors* that are needed to execute all the steps in between. The generic workflow below is expressed in a structured format in which keywords are expressed in bold and roles are expressed in italics.

The keywords “may”, “may be” and “or” indicate variation points across companies or within a company across projects. For example, the translation tool is defined for several companies, while it is not specified for others. The tools for sending a translation request, however, can vary from project to project. The keyword “does” indicates a required step, sometimes limited to specific cases by an “if” condition. Instantiations of this generalized workflow may contain less steps (non-required steps can be left out) or be more severely constrained; steps that are optional in general may be required for a specific company.

*Project manager is internal*

*Language engineer is internal or unspecified*

*Translator is freelancer*

*Auditor is freelancer or internal*

*Supervisor is internal or unspecified*

*scoring may be required*

*preparation may be required*

*translation tool is SDL Trados Studio or in-house developed*

*translation tool or unspecified*

1. *Client does* send Translation request **using** email **or** telephone **or** specific tool
2. *Auditor or Language engineer may* prepare **one or more** translation packages
  - a. **may** optimize content
  - b. **may** split content

**per translation package:**

3. *Project manager does* select potential *Translator*
4. *Project manager does* contact *Translator* **using** email **or** dedicated tool
  - a. **does** contact *Auditor* if *Auditor is freelancer*
5. *Translator does* confirm **or** deny **using** email **or** specific tool

**if *Translator* did confirm:**

6. *Project manager does* send translation package **to** *Translator*
7. *Translator does* translate **using** translation tool
8. *Translator does* send translated document **to** *Project manager*
9. *Project manager does* send translated document **to** *Auditor*
10. *Auditor does* verify translated document
11. *Auditor may* send updated translated document **to** *Translator*
12. *Auditor or Project manager does* score *Translator* **if** scoring **is** required
13. *Supervisor does* perform quality check **if** *Supervisor is internal*

14. *Project manager does* merge translated documents **if** content **was** split
15. *Project manager does* send translated document **to** *Client*
16. *Client may* send feedback **to** *Project manager*

While the workflow is expressed in a generalized form, not all steps are explicitly represented; e.g. transfer of the document between different roles within the company is not explicitly mentioned. Other information that was left out includes the criteria for selecting and scoring a translator as these were mostly or completely consistent; respectively availability and competences, and EN 15038.

The terminology was made consistent across companies (e.g. the *Auditor* role was not mentioned in one interview). Two roles, *language engineer* and *supervisor*, were only mentioned in one company. The exact content of the translation package was not taken into account in this workflow as it was not mentioned explicitly in all interviews. The generalized workflow uses a single term, send, to indicate a transfer of a document to another person involved in the workflow regardless of the specific modality that is used, which may include sharing through a web services or web portal.

The translation agencies that handle translation assignments on a daily basis have four steps in common: *selection of translator(s)*, *translator accepts or rejects assignment*, *translator translates document*, *translated documents are revised by auditor or translator*. We need to mention that the revision<sup>43</sup> itself is different for these companies. One company selects a freelance translator to make the revision, while two other companies have in-house auditors to fulfil this task. For all translation agencies considered here, we can also conclude that *request of translation assignment* and *sending translation to customer* are steps in the general workflow.

There are also differences in the way that the translation agencies get the translation requests. There is one common way, through either email or telephone. However, one translation agency also disposes of an online interface, whereas another company receives requests through SDL Worldserver as well. They all send the translated documents to the customer by email despite the availability of an online interface for one of the companies.

## Conclusions

Almost all companies that were interviewed are working with freelance translators, whereas the type of auditors they are working with varies from company to company. Some companies prefer to have in-house auditors while others prefer to select a second freelance translator to revise the translation.

The companies have different prospects concerning the future. Some companies are looking forward to having a cloud-based solution, while others fear privacy issues.

Interoperability or the files exchanged between different translation environments remains an issue. Further, there is a need for flexible user interface-designs with customization options that allow users to adapt the tools to their own workflows. Live preview or WYSIWYG is a desirable feature within the translation editor.

All interviewed companies prefer assigning one translator per job, rather than allowing multiple translators and reviewers work on the same document at the same time, in order to minimize inconsistencies in the final document. With regards to the relationship between clients and translation vendors, clients are not involved in the translation process and they hardly provide any feedback after the translation has been delivered.

## Contextual inquiries

Insights in the actual work practices of professional translators were obtained through contextual inquiries --- a user research technique that combines observations and

---

<sup>43</sup> Making of changes to a translation, either by the translator (“self-revision”) or someone else (“other-revision”). (Source: Pym, 2011)

interviews. (Beyer & Holtzblatt, 1998) (Karen Holtzblatt, Jessamyn B. Wendell, and Shelly Wood, 2004) Observation of the workflow of language professionals can contribute to the identification of specific needs, translation tools that are used, the way translation tools are personalized and difficulties that are faced during the translation assignment can be detected. For the recruitment of those language professionals, we contacted, besides the partners of the *IAC*<sup>44</sup>, some freelance translators and other companies where translation is one of their daily activities.

The goal of the contextual inquiries reported on in this deliverable is to reveal opportunities for improving the efficiency of professional translators and different workflows of professional translators.

First, we discuss the participants and a summary of the key observations related to their profile and how they work. .Next, we present observations about TEnTs and how they were used. We conclude the section with the definition of two personas, hypothetical characters based on the user research (contextual inquiry, but also the other parts to the research presented in this deliverable).

## Participants

We conducted nine contextual inquiries that involved seven translators, one supervisor and one team that provide captions on broadcast series were observed. We decided to observe language professionals with different profiles so that we could compare potential differences in workflows.

The participants were asked to sign an informed consent form (Lazar, Feng, & Hochheiser, 2010) before the observations. The participants thus gave permission to be observed.

In the discussion of the results below, any reference to the nine participants will be made using an anonymous name, such as P1 or P2. The profession, education, experience, native and active languages, domain experience, personal computer (PC) setup, software, file types, training and profile will be specified for each participant.

A workflow for each participant is defined based on the observation.

Each workflow lists all the steps that were performed from the request for translation until the delivery of the translated documents back to the client.

---

<sup>44</sup> Industry Advisory Commission

### Overview

Participant	Profession	Linguistic experience (in years)	Active languages (native language)	Software experience (currently used TEnT)
P1	Project manager (30%) In-house translator (70%)	10+	<u>Dutch</u> , French, English, German	<u>SDL Trados Studio 2009 &amp; 2014</u> MS Office, Internet Explorer
P2	In-house translator	5+	<u>Dutch</u> , French, English, German	<u>SDL Trados Studio 2009 &amp; 2014</u> MS Office, Internet Explorer
P3	Freelance translator	20+	<u>Dutch</u> , English	<u>SDL Trados Studio 2007</u> (only pre-translation), <u>Déjà Vu X3</u> ABBYY FineReader <sup>45</sup> , MS Excel Adobe PDF, TextPad <sup>46</sup> Internet Explorer (Google Search, Van Dale)
P4	Freelance translator	18	<u>French</u> , English	<u>SDL Trados Studio 2011</u> WordFast, ForeignDesk <sup>47</sup> , Microsoft LocStudio, LionBridge <sup>48</sup> , Dropbox <sup>49</sup> , MS Office, Internet Explorer (Google Search, IATE)
P5	Freelance translator	10	<u>Dutch</u> , English, French, German, Italian	<u>In-house developed TEnT (provided by client)</u> , <u>SDL Trados Studio</u> <u>SDL Passolo</u> <sup>50</sup> , Microsoft LocStudio MS Word, Internet Explorer
P6	Linguistic Services Supervisor	10+	<u>German</u> , French, English, Spanish, Italian	Translation Quality Tool (TQT, in-house tool) Xbench, Microsoft Excel, Google Chrome
P7	Freelance translator	20+	<u>Dutch</u> , English	Softel Swift <sup>51</sup> , Citrix XenApp Microsoft Word, Adobe PDF Google Chrome
P8	Provider of closed captions	Unknown	<u>Dutch</u>	Softel Swift, Dragon Naturally Speaking <sup>52</sup>
P9	In-house translator	10	<u>Dutch</u> , English	<u>SDL Trados Studio</u> , TEnT Integration tool Internet Explorer

Table 5 Overview of participants of contextual inquiries

### Participant Description

This section provides more information about the contextual inquiries that were performed. The tables contain profession-related information of the observed participants and the workflow that is derived from the contextual inquiry. In addition, each table contains a profile where specific characteristics of the participant are mapped on several scales.

<sup>45</sup> <http://finereader.abbyy.com/>

<sup>46</sup> <https://www.textpad.com/>

<sup>47</sup> <http://www.foreigndesk.net>

<sup>48</sup> <http://www.lionbridge.com/>

<sup>49</sup> <https://www.dropbox.com/>

<sup>50</sup> <http://www.sdl.com/exc/language/software-localization/passolo/>

<sup>51</sup> [http://www.grassvalley.com/products/swift\\_create](http://www.grassvalley.com/products/swift_create)

<sup>52</sup> <http://www.nuance.com/dragon/index.htm>

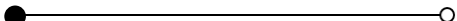


### Scale description

Different characteristics detected during the contextual inquiries are used to specify the profile of the participant. In order to compare and align the different profiles, we describe each profile using scales for specific characteristics of each participant. The scales are based on a 5-point scale, except the scale that indicates the number of languages, which contains two options. The scales used are listed below.


### Interaction technique

This scale defines how the participant navigates through and interacts with the TEnT. It can be either the mouse or the keyboard or a combination of both.

Mouse  Keyboard


### Media

This scale indicates how often the participant uses paper or digital documents while translating or revising.

Paper  Digital


### Software knowledge

This scale specifies if a participant is a novice, intermediate or advanced user of the used TEnT. The definition of this factor is based on general knowledge of the features, the customization and the structuring of work practices.

Limited  Extensive

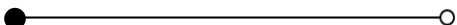
### Training

This scale defines how much training the participant received to learn how to use the TEnT.

No training  Training


### Support

This scale indicates if the company or colleagues support the translator in using the TEnT. By support, we mean that someone provides help or tips and tricks on how to work with the TEnT.

No support  Support

### Active languages

This scale indicates the amount of languages the participant is working with and shows “two languages” or “more than two languages”.

Two languages  More than two languages

### Customization

Customization defines to which degree the participant has changed some settings in the TEnT to adapt it to his individual needs, e.g. changes default keyboard shortcuts, changes settings regarding source and target language, terminology.

No customization ●————○ Customization

### Professional environment

This scale indicates if the participant is working in-house or freelance or a combination of both.

In-house ●————○ Freelance

### Participant P1

During the contextual inquiries, P1 was translating a French business related Word document containing a small number of pages into Dutch. Besides this translation job, we were also shown a partial translation of a PowerPoint file. P1 also assigns translation requests to the other translators in the company. P1 is currently using SDL Trados 2014.

Participant P1 uses some keyboard shortcuts to navigate through the editor of the TEnT while translating. Even though he is using SDL Trados Studio, he still uses a lot paper documents to support his work. He prints all the slides on paper when a PowerPoint presentation needs to be translated. He also prints the translated document on paper to read the translation and make some corrections where needed. His knowledge about the TEnT is very limited, the one day training he got from a trainer of SDL Trados Studio. P1 translates from French to Dutch in 95% of his translation assignments. Besides that, he sometimes needs to translate from or to English and German. P1 didn't customize the tool, he uses some of the default keyboard shortcuts and kept the default matching values.

<b>Education</b>	Master Linguistics
<b>Domain experience</b>	Business, Commerce, Management and Marketing Insurance
<b>PC setup</b>	Single screen, desktop
<b>File types</b>	Translation file (*.xliff), Initial file type (*.docx, *.pptx, *.xlsx, *.pdf)
<b>Training</b>	One day training from SDL Trados Studio in own company, other features self-taught
<b>Workflow</b>	<ul style="list-style-type: none"> <li>● P1 receives translation request through email</li> <li>● P1 saves documents from email on local drive</li> <li>● P1 starts SDL Trados Studio 2014               <ul style="list-style-type: none"> <li>○ P1 selects 'Translate Single Document' in menu bar</li> <li>○ P1 selects source and target language</li> <li>○ P1 selects translation memory (stored on server)</li> </ul> </li> </ul>

	<ul style="list-style-type: none"> <li>• P1 starts translation <ul style="list-style-type: none"> <li>◦ P1 types translation in target column</li> <li>◦ SDL Trados Studio automatically fills out fuzzy matches (&gt; 70%)</li> <li>◦ P1 presses 'ctrl + enter' to confirm segment and go to next segment</li> <li>◦ P1 consults Termias (in-house termbase) to check for translation of word</li> <li>◦ P1 finishes translation</li> </ul> </li> <li>• P1 saves translation (.xliff extension)</li> <li>• P1 saves translation as initial document type</li> <li>• P1 revises his own translation on paper or in software of initial document type</li> <li>• P1 makes changes in SDL Trados Studio 2014 based on remarks made when revising</li> <li>• P1 sends translation in initial document type to final recipient(s)</li> <li>• P1 stores email request in folder with finished translations</li> </ul>			
Profile	Interaction technique	Mouse		Keyboard
	Media	Paper		Digital
	Software knowledge	Limited		Extensive
	Training	No training		Training
	Support	No support		Support
	Active languages	Two languages		More than two languages
	Customization	No customization		Customization
	Professional environment	In-house		Freelance



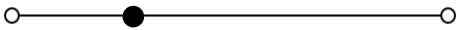
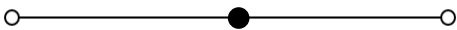
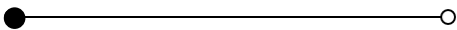


*Table 6 Overview of contextual inquiry of P1*

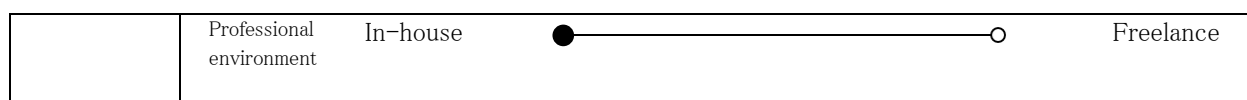
## Participant P2

P2 is working in the same company as P1 and was translating a business-related Word document containing a small number of pages. P2 is currently using SDL Trados 2014. P1 and P2 sometimes work together when translating bigger projects with a strict deadline.

Participant P2 uses some of the default keyboard shortcuts to navigate through the editor of the TEnT while translating. P2 prints the translation on paper to read it and correct some mistakes. She received a one day training in SDL Trados Studio. Her knowledge of the TEnT is restricted to the basic features, i.e. translating using the editor, adding the required files to the project and limited customization of project settings. An example of this customization is that P2 sets a default source and target language, and a default translation memory. This removes these steps during project creation for almost all

translation assignments done by P2 (from French to Dutch). Less than 5% is from or to English or German where Dutch is the source or target language in these cases.

<b>Education</b>	Master Germanic Philology			
<b>Domain experience</b>	Business, Commerce, Management and Marketing Insurance			
<b>PC setup</b>	Dual screen, desktop			
<b>File types</b>	Translation file (*.xliff), Initial file type (*.docx, *.pptx, *.xlsx, *.pdf)			
<b>Training</b>	One day training from SDL Trados Studio in own company, other features self-taught			
<b>Workflow</b>	<ul style="list-style-type: none"> <li>• P2 receives translation request through email</li> <li>• P2 saves documents from email on local drive</li> <li>• P2 opens source document on left screen</li> <li>• P2 starts SDL Trados 2014 on right screen               <ul style="list-style-type: none"> <li>◦ P2 selects 'Translate Single Document' in menu bar</li> </ul> </li> <li>• P2 starts translation               <ul style="list-style-type: none"> <li>◦ P2 types translation in target column</li> <li>◦ SDL Trados automatically fills out fuzzy matches (&gt; 70%)</li> <li>◦ P2 presses 'ctrl + enter' to confirm segment and go to next segment</li> <li>◦ P2 consults Termias (in-house termbase) to check for translation of word</li> <li>◦ P2 turns off tags in SDL Trados Studio if they are overwhelming in segment</li> <li>◦ P2 finishes translation</li> </ul> </li> <li>• P2 saves translation (.xliff extension)</li> <li>• P2 saves translation as initial document type</li> <li>• P2 revises his own translation on paper or in software of initial document type (e.g. MS Word, MS Excel, MS PowerPoint)</li> <li>• P2 makes changes in SDL Trados Studio 2014 based on remarks made when revising</li> <li>• P2 sends translation in initial document type to final recipient(s)</li> <li>• P2 stores email request in folder with finished translations</li> </ul>			
<b>Profile</b>	Interaction technique	Mouse		Keyboard
	Media	Paper		Digital
	Software knowledge	Limited		Extensive
	Training	No training		Training
	Support	No support		Support
	Active languages	Two languages		More than two languages
	Customization	No customization		Customization



*Table 7 Overview of contextual inquiry of P2*

### Participant P3

P3 was working on the translation of an extensive software manual. The project contained 4275 files comprising 1 400 000 words that needed to be translated from English into Dutch. The project consisted of merely XML files. On big projects P3 works very structured: he always creates a fixed set of folders where he saves the corresponding documents. He also spends several days on project planning. P3 uses SDL Trados 2007 to pre-translate the files and Déjà Vu X3 to translate the files.

Participant P3 uses many keyboard shortcuts to navigate through the editor of the TEnT. P3 has customized a lot of features to make it as easy as possible for him to work with the TEnT. He sets a default master translation memory and a default term base, he disables features, e.g. AutoWrite, that distract him while translating, he adds new terms to the term base while translating to keep the term base up to date in every stage of the translation process.

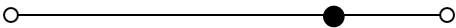
He only prints an overview of the files that needs to be translated for a specific translation assignment and the regarding numbers (i.e. number of files, word count, 100% matches count ...). He also makes use of paper dictionaries, e.g. ICT dictionary, to search for a certain translation if online dictionaries like e.g. Van Dale doesn't provide a translation.


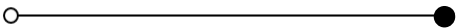
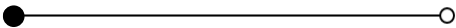

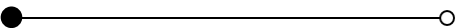


P3 is an advanced user of SDL Trados and Déjà Vu and he learnt to work with these technologies by himself, without any support from his regular clients. He regularly attends translators' conferences to keep up to date with the developments in the translation technology field.

Besides translation, P3 is passionate about terminology management and has recently attended the Terminology Summer School<sup>53</sup> in Vienna to prepare for the Terminology Manager Basic Certificate.

<b>Education</b>	Linguistic background
<b>Domain experience</b>	IT, Data processing, Computing, Electronics, Control Engineering (User interfaces and manuals of software)  Medicine and Pharmaceuticals  Applied Social Sciences (Sociology, Psychology, Statistics, Human Geography)
<b>PC setup</b>	Single screen, desktop

<sup>53</sup> Terminology Summer School organized by TermNet.

<b>File types</b>	Translation file (*.ttx), Initial file type (*.xml, *.docx, *.pptx, *.xlsx, *.pdf)
<b>Training</b>	Self-taught
<b>Workflow</b>	<ul style="list-style-type: none"> <li>• P3 downloads project via FTP</li> <li>• P3 saves documents in a fixed structure in specific folders <ul style="list-style-type: none"> <li>◦ E.g. 01 source, 02 ..., 03 ..., 04 Translation, 05 Log, 06 Finance, 07 Review, ...</li> <li>◦ P3 saves back up on external hard disk</li> </ul> </li> <li>• P3 reads style instructions provided by client</li> <li>• P3 does pre-translation in SDL Trados Studio 2007 <ul style="list-style-type: none"> <li>◦ P3 starts SDL Trados Studio 2007</li> <li>◦ P3 opens XML source document in SDL Trados Studio</li> <li>◦ P3 does a simple quality check to check if all XML tags are in right order</li> <li>◦ P3 saves</li> </ul> </li> <li>• P3 opens log file of project in TextPad to check number of words, size of project, matches</li> <li>• P3 calculates invoice in MS Excel based on numbers from log file</li> <li>• P3 opens CSV file containing files that needs to be translated in MS Excel <ul style="list-style-type: none"> <li>◦ P3 sorts columns on word count (descending order)</li> </ul> </li> <li>• P3 makes project planning (files/day)</li> <li>• P3 opens Déjà Vu <ul style="list-style-type: none"> <li>◦ P3 creates new project</li> <li>◦ P3 selects source and target language</li> <li>◦ P3 selects client and domain</li> <li>◦ P3 opens file in Déjà Vu</li> <li>◦ P3 changes configurations <ul style="list-style-type: none"> <li>▪ P3 disables AutoWrite</li> </ul> </li> <li>◦ P3 starts translating text by translating segments in right order (top-down) <ul style="list-style-type: none"> <li>▪ P3 copies source text to target text and translates word per word in segment</li> <li>▪ When P3 translates a word that occurs a lot in the text, he already translates the same ones by using find-replace and selecting word-only</li> <li>▪ P3 uses AutoSearch to check for translation in termbase or context</li> <li>▪ P3 scans translation memory for translation of certain word</li> <li>▪ P3 uses AutoSave after translation of segment by pressing ctrl + arrow down to confirm segment and save it</li> <li>▪ P3 uses paper ICT dictionary to translate the word 'overflow'</li> <li>▪ P3 adds term from source text manually to termbase by right clicking on mouse - add term <ul style="list-style-type: none"> <li>• P3 specifies word type (noun, verb, ...), gender, number, subject, client and context for both source and target text</li> </ul> </li> <li>▪ P3 consults online Q&amp;A document if P3 doesn't know context of word to ask the client for more clarity</li> <li>▪ P3 doesn't know the article of 'prefix' <ul style="list-style-type: none"> <li>• P3 consults 'Groene Boekje' to see if it is 'de' or 'het'</li> </ul> </li> <li>▪ P3 doesn't know if 'templated' in source text is an existing word <ul style="list-style-type: none"> <li>• P3 consults online Van Dale to check if it exists</li> </ul> </li> </ul> </li> <li>◦ P3 finishes translation <ul style="list-style-type: none"> <li>▪ P3 saves document as initial document type</li> <li>▪ P3 checks layout of final document <ul style="list-style-type: none"> <li>• P3 corrects layout if required changes are limited</li> </ul> </li> </ul> </li> <li>◦ P3 has translated all files and sends result via FTP to client</li> </ul> </li> </ul>
<b>Profile</b>	Interaction technique      Mouse  Keyboard

Media	Paper		Digital
Software knowledge	Limited		Extensive
Training	No training		Training
Support	No support		Support
Active languages	Two languages		More than two languages
Customization	No customization		Customization
Professional environment	In-house		Freelance

*Table 8 Overview of contextual inquiry of P3*

### Participant P4

P4 is a technical translator with the English/French language combinations. P4 is mainly working for one translation agency that prefers their freelance translators to use SDL Trados Studio.

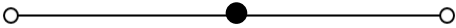
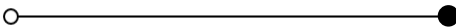
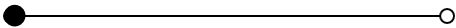

During the contextual inquiry, P4 was translating a document, which described a defect of an electronic device, from English into French. The quality of the source language was bad, which made it more difficult to translate the manual.

Participant P4 uses a few keyboard shortcuts while translating, e.g. a shortcut to confirm a translated segment. When searching the termbase or the translation memory, P4 uses the mouse. P4 uses only digital devices to fulfil the translation job.

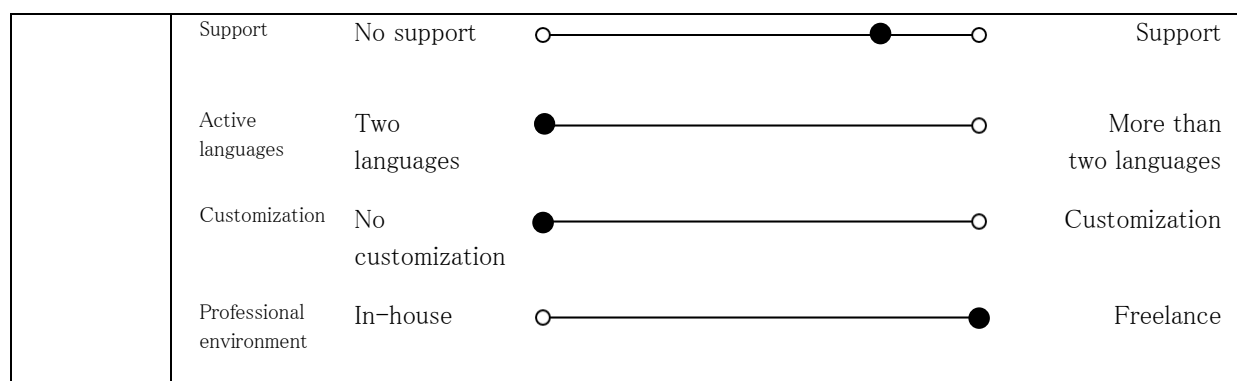
P4's knowledge of SDL Trados is limited to the usage of the editor. She didn't customize the tool. The limited knowledge may be a result of the fact that she didn't receive any training on the tool. However, she receives technical support from her client when needed. The version of SDL Trados Studio she is working with is outdated, which may cause problems sometimes. For example, P4 was not able to view and edit terms in the MultiTerm directly from the editor interface because of Java-related errors<sup>54</sup>.

<b>Education</b>	Linguistic background
<b>Domain experience</b>	IT, Data processing, Computing, Electronics, Control Engineering (Manuals of electronic devices)

<sup>54</sup> <http://kb.sdl.com/kb/?ArticleId=4956&source=article&c=12&cid=23#tab:homeTab:crumb:7:artId:4956>,

	Medicine and Pharmaceuticals (Manuals of medical devices) Business, Commerce, Management and Marketing (limited)			
<b>PC setup</b>	Dual screen, desktop and laptop (laptop screen is duplicate of desktop screen)			
<b>File types</b>	Translation file (*.xliff), Initial file type (*.docx, *.pptx, *.xlsx, *.pdf)			
<b>Training</b>	Self-taught			
<b>Workflow</b>	<ul style="list-style-type: none"> <li>• P4 downloads package (documents) for translation assignment through FTP server or Dropbox <ul style="list-style-type: none"> <li>○ Small project (one file) is sent via e-mail</li> <li>○ Package already contains .xliff file</li> <li>○ Pre-translation already done</li> </ul> </li> <li>• P4 opens SDL Trados Studio 2011 <ul style="list-style-type: none"> <li>○ P4 opens .xliff file in SDL Trados Studio 2011</li> <li>○ P4 starts translating each segment <ul style="list-style-type: none"> <li>▪ P4 switches often between source document and SDL Trados Studio 2011</li> <li>▪ P4 copies source text in target text segment if no translation is suggested <ul style="list-style-type: none"> <li>• P4 translates the source text word by word to make sure that he doesn't forget a part of the segment</li> </ul> </li> <li>▪ P4 uses different sources if P4 doesn't know translation or meaning of word <ul style="list-style-type: none"> <li>• French online dictionary, IATE, Terminology Research Engine System, Alchemy, Wikipedia, Electropedia, Linguee</li> </ul> </li> <li>▪ P4 consults online Q&amp;A file if P4 can't find translation of a word</li> <li>▪ P4 hesitates about translation of 'hydro unit', P4 searches two possible translations on Google to see which one most occurs in the search results</li> </ul> </li> <li>○ P4 finishes translation <ul style="list-style-type: none"> <li>▪ P4 confirm every translated segment one by one with a keyboard shortcut</li> <li>▪ P4 executes spell check by pressing F7 in SDL Trados Studio</li> <li>▪ P4 copies target text to MS Word to execute a spell (and grammar) check, because SDL Trados Studio doesn't check on grammar</li> </ul> </li> <li>○ P4 saves document with the extension .xliff</li> </ul> </li> <li>• P4 updates own project management sheet in MS Excel</li> <li>• P4 makes a zip of translated documents</li> <li>• P4 uploads zip via FTP</li> </ul>			
<b>Profile</b>	Interaction technique	Mouse		Keyboard
	Media	Paper		Digital
	Software knowledge	Limited		Extensive
	Training	No training		Training





*Table 9 Overview of contextual inquiry of P4*

### Participant P5



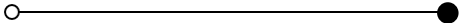
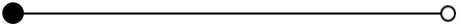

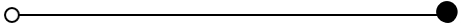


During the contextual inquiry, P5 was revising a document (source language was English, and target Dutch) but he gets translation requests in 70% of the cases. He usually works for the same translation agency.

He has a structured way of checking specific errors in the test, e.g. checking for double spaces, capital letters, punctuation, and spelling and grammatical errors.

Participant P5 uses a lot of keyboard shortcuts to navigate through the editor of the TEnT while translating. He customized some keyboard shortcuts to be more convenient. P5 uses only his personal computer to fulfil the translation or revision assignment.

P5 knows the TEnT he is working with very well, it is less complex than e.g. SDL Trados Studio. He also has experience with SDL Trados Studio, but the complexity is one the reasons he is not using it anymore. P5 did not receive any training, but the support provided by his client is very good. Several times a year, all translators working with that agency are notified per e-mail of software updates. Last but not least, P5 remarks that translators have the possibility to suggest changes or improvements.

<b>Education</b>	Linguistic background
<b>Domain experience</b>	Accounting, finance, banking, economics Environment/Environmental engineering and sciences/Water Business, Commerce, Management and Marketing
<b>PC setup</b>	Single screen, desktop
<b>File types</b>	Translation file (*.lwo)
<b>Training</b>	Self-taught + support by client
<b>Workflow</b>	<ul style="list-style-type: none"> <li>• P5 downloads documents from online portal of client he is working for</li> <li>• P5 opens in-house developed TEnT</li> <li>• P5 opens translated document in in-house developed TEnT</li> <li>• P5 starts revision <ul style="list-style-type: none"> <li>○ P5 uses CTRL + F to search for double spaces</li> </ul> </li> </ul>

	<ul style="list-style-type: none"> <li>○ P5 performs a spellcheck <ul style="list-style-type: none"> <li>▪ Spellcheck gives list of errors below the editor window</li> <li>▪ P5 adds exceptions to spellcheck list based on false positive errors that were shown (e.g. Dhr.)</li> </ul> </li> <li>○ P5 copies target text to MS Word to perform a grammar check</li> <li>○ P5 performs a consistency check (provided in in-house developed TEnT)</li> <li>○ P5 scrolls fast through target text to check for specific colours (that indicate certain errors, such as non-translated words)</li> <li>○ P5 places bookmark sometimes in text (to switch back to it afterwards)</li> <li>○ P5 corrects mistakes in target text</li> <li>• P5 finishes revision</li> <li>• P5 uploads revised documents to online portal</li> </ul>			
<b>Profile</b>	Interaction technique	Mouse		Keyboard
	Media	Paper		Digital
	Software knowledge	Limited		Extensive
	Training	No training		Training
	Support	No support		Support
	Active languages	Two languages		More than two languages
	Customization	No customization		Customization
	Professional environment	In-house		Freelance

*Table 10 Overview of contextual inquiry of P5*

## Participant P6

P6 was performing a final QA check on an automotive manual during the contextual inquiry. The manual had already been translated and revised. P6 was using a proprietary stand-alone QA tool to revise the document.

Participant P6 uses a lot of keyboard shortcuts to navigate through the editor while translating. He also customized some keyboard shortcuts. P6 only uses his personal computer to fulfil the translation or revision assignment. P6 knows the TEnT he is working with very well and customizes a lot of features, such as the avoidance of false positives through regular expressions.

He knows SDL Trados Studio, but he does not use it since they had a proprietary TEnT-tool. The agency he was working for developed their own translation tool to meet the specific needs of the Quality Assessment assignments. P6 was an in-house translator when

the current TEnT was under development and got to know all its features through the development team. Currently there are no more developers available, and thus new users are only able to learn the tool through the aid of colleagues or the manual written by P6.

P6 sometimes has to work with languages that he does not really know, such as Spanish. This does not pose a problem since he has a linguistic background and only has to perform Quality Assessment of translated documents. P6 is a heavy tinkerer of the system and he will always try to find new and easier ways to complete the assignments.

<b>Education</b>	Linguistic background
<b>Domain experience</b>	Business, Commerce, Management and Marketing Motor vehicles, Road/Rail transport
<b>PC setup</b>	Dual screen, desktop
<b>File types</b>	Translation file (*.ttx), Translation feedback (*.xlsx)
<b>Training</b>	Self-taught + explanation by in-house developers
<b>Workflow</b>	<ul style="list-style-type: none"> <li>• P6 receives incoming assignment via e-mail</li> <li>• P6 creates folder creation for assignment on shared drive <ul style="list-style-type: none"> <li>◦ P6 creates subfolders per destination language</li> <li>◦ P6 copies files to destination folder</li> </ul> </li> <li>• P6 opens a remote desktop connection to a virtual environment with an XP OS</li> <li>• P6 opens TQT <ul style="list-style-type: none"> <li>◦ P6 selects the language pairs to load the translation profile</li> <li>◦ P6 opens the ttx file on the shared drive</li> </ul> </li> <li>• P6 opens XBench <ul style="list-style-type: none"> <li>◦ P6 loads the ttx file in XBench</li> </ul> </li> <li>• P6 reviews the translation line per line in TQT</li> <li>• If a wrong translation occurs <ul style="list-style-type: none"> <li>◦ P6 identifies and corrects the mistake</li> </ul> </li> <li>• If the error given by TQT is unclear <ul style="list-style-type: none"> <li>◦ P6 checks the code visible in the statusbox to identify issue</li> </ul> </li> <li>• If the error is still unclear: line is copied to XBench to retrieve context</li> <li>• If a false positive occurs, P6 right clicks on the line and selects “Ignore all” <ul style="list-style-type: none"> <li>◦ Recurring false positive: P6 adds regex to preferences to ignore and reloads the profile settings</li> </ul> </li> <li>• When all lines in current screen are corrected, P6 reloads the translation</li> <li>• If the source text is unclear or contains an error <ul style="list-style-type: none"> <li>◦ P6 e-mails the creators of the document in Japan to clarify</li> </ul> </li> <li>• If this is a large project <ul style="list-style-type: none"> <li>◦ P6 applies a change tracker to the translated files</li> <li>◦ P6 imports the results in a predefined Excel sheet</li> <li>◦ Excel calculates statistics (amount of changes, severity level)</li> <li>◦ P6 cross-references the statistics with reviewer report</li> <li>◦ P6 adds the statistics to the end results</li> </ul> </li> <li>• If the QA assignment is done finished <ul style="list-style-type: none"> <li>◦ P6 sends the QA results via e-mail to close the assignment</li> <li>◦ P6 sends the translation feedback to the translator via e-mail</li> <li>◦ P6 sends the source text comments to the creators of the document in Japan</li> </ul> </li> </ul>

Profile	Interaction technique	Mouse		Keyboard
	Media	Paper		Digital
	Software knowledge	Limited		Extensive
	Training	No training		Training
	Support	No support		Support
	Active languages	Two languages		More than two languages
	Customization	No customization		Customization
	Professional environment	In-house		Freelance

*Table 11 Overview of contextual inquiry of P6*

### Participant P7

P7 was providing Dutch captions on English TV series during the contextual inquiry. She uses the Swift software to provide the subtitles. P7 works always for the same client.

During translation and captioning, P7 uses the keyboard all the time. For other tasks, P7 relies on the mouse. P7 only uses her personal computer to fulfil the translation assignment. P7 only knows the basics of the tool she uses and does not invest time in exploring the other features because she believes the basics are good enough. She received a basic training from one of her colleagues and learned the rest by herself. There is no support available for the tool other than help from colleagues. P7 does not know any other TEnTs.

<b>Education</b>	Linguistic background
<b>Domain experience</b>	Other
<b>PC setup</b>	Single screen, laptop
<b>File types</b>	Caption File (unknown ext.), Video file (*.wmv; *.mp4), Video script (*.doc; *.pdf)
<b>Training</b>	Self-taught + explanation by colleagues
<b>Workflow</b>	<ul style="list-style-type: none"> <li>• P7 receives incoming assignment via the Workflow website</li> <li>• P7 opens Swift             <ul style="list-style-type: none"> <li>◦ P7 imports the assignment by reference number</li> <li>◦ P7 imports the video file located on shared drive</li> </ul> </li> <li>• P7 starts playback of video             <ul style="list-style-type: none"> <li>◦ When a dialog starts, P7 indicates the in-time using a numpad key</li> </ul> </li> </ul>

	<ul style="list-style-type: none"> <li>○ P7 listens to dialog <ul style="list-style-type: none"> <li>▪ If the dialog is unclear, P7 refers to the script</li> </ul> </li> <li>○ When the dialog ends, P7 indicates the out-time using a numpad key</li> <li>• P7 pauses the video and enters the translation of the dialog</li> <li>• If the translation is unknown <ul style="list-style-type: none"> <li>○ P7 searches for the translation using Google Search, Google Translate, online Van Dale and/or a digital XenApp dictionary</li> </ul> </li> <li>• If the caption is too long (CPS count too high) <ul style="list-style-type: none"> <li>○ Or P7 rephrases the translation</li> <li>○ Or P7 redefines the out-time</li> </ul> </li> <li>• If the translation is uncommon <ul style="list-style-type: none"> <li>○ P7 adds a note to the caption to explain the translation to the responsible</li> </ul> </li> <li>• If the translation has been done before and P7 wants to find it <ul style="list-style-type: none"> <li>○ P7 searches through the captions (Ctrl+F)</li> </ul> </li> <li>• If P7 jumps to the start of the caption with a numpad key (Del) <ul style="list-style-type: none"> <li>○ If the in-time is not perfect, P7 adjusts the in-time using F5 and F6</li> </ul> </li> <li>• P7 confirms the caption</li> <li>• If the captioning is complete <ul style="list-style-type: none"> <li>○ P7 checks the Diagnostic Window and corrects each error</li> </ul> </li> <li>• If the captioning is complete and there are no more errors <ul style="list-style-type: none"> <li>○ P7 saves the result on the local disk</li> <li>○ P7 uploads the result to the Workflow website</li> </ul> </li> </ul>			
<b>Profile</b>	Interaction technique	Mouse		Keyboard
	Media	Paper		Digital
	Software knowledge	Limited		Extensive
	Training	No training		Training
	Support	No support		Support
	Active languages	Two languages		More than two languages
	Customization	No customization		Customization
	Professional environment	In-house		Freelance

Table 12 Overview of contextual inquiry of P7

### Participant P8a, P8b, P8c

P8a, P8b and P8c are a team of three people who provide Dutch captions on Dutch TV live broadcasts. They were providing captions of a political TV show during the contextual inquiry

For this purpose, they were using speech recognition software Nuance Dragon Naturally Speaking (DNS) in combination with the subtitling software Softel Swift.

P8a “respeaks” the content of the video by using Nuance Dragon Naturally Speaking and the text is automatically inserted into Swift. P8b corrects the text that was inserted in Swift and P8c performs a last check and then releases the captions for broadcasting the TV show.

During captioning, P8 uses the keyboard all the time. For other tasks, P8 relies on the mouse. P8 only uses his personal computer to fulfil the translation assignment. P8 only knows the basics of the tool she uses and does not spend extra time exploring all the features of the software. She received a short training by one of her colleagues and had to learn the rest by herself. P8 does not know any other TEnTs. There is no support available for the tool other than help from colleagues. P8 does not translate documents, but only provides captioning for live broadcasts.

<b>Education</b>	Linguistic background
<b>Domain experience</b>	Other
<b>PC setup</b>	Dual screen, desktop
<b>File types</b>	Caption file (unknown ext.)
<b>Training</b>	Self-taught + training from colleagues
<b>Workflow</b>	<p>P8a:</p> <ul style="list-style-type: none"> <li>• P8a opens Swift <ul style="list-style-type: none"> <li>◦ Live program is playing on second screen</li> </ul> </li> <li>• When a dialog starts, P8a indicates the start of the caption (Numpad key)</li> <li>• P8a listens to the dialog <ul style="list-style-type: none"> <li>◦ P8a respects the dialog in Dragon Naturally Speaking</li> </ul> </li> <li>• If there is a punctuation mark <ul style="list-style-type: none"> <li>◦ P8a types the punctuation mark</li> </ul> </li> <li>• If there is a difficult word in terms of speech-to-text <ul style="list-style-type: none"> <li>◦ P8a types the word</li> </ul> </li> <li>• If there is an error in recognition and there is still time available <ul style="list-style-type: none"> <li>◦ P8a corrects the error</li> </ul> </li> <li>• If the caption is too long <ul style="list-style-type: none"> <li>◦ P8a types “...”, sets out-time and starts new caption (in-time)</li> </ul> </li> <li>• P8a confirms caption</li> </ul> <p>P8b:</p> <ul style="list-style-type: none"> <li>• P8b opens Swift <ul style="list-style-type: none"> <li>◦ Live program is playing on second screen (with delay)</li> </ul> </li> <li>• P8b reviews caption made by P8a</li> <li>• If there is an error <ul style="list-style-type: none"> <li>◦ P8b corrects caption</li> </ul> </li> <li>• P8b confirms caption</li> </ul> <p>P8c:</p>

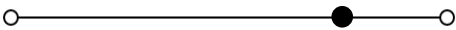


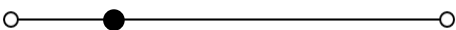

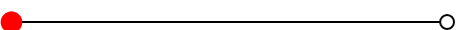
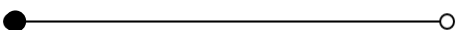

	<ul style="list-style-type: none"> <li>• P8c opens Swift <ul style="list-style-type: none"> <li>◦ Live program is playing on second screen (40 minutes)</li> </ul> </li> <li>• P8c reviews captions made by P8a, reviewed by P8b</li> <li>• If there is an error <ul style="list-style-type: none"> <li>◦ P8c corrects caption</li> </ul> </li> <li>• P8c confirms caption <ul style="list-style-type: none"> <li>◦ Displayed on live broadcast</li> </ul> </li> </ul>			
Profile	Interaction technique	Mouse		Keyboard
	Media	Paper		Digital
	Software knowledge	Limited		Extensive
	Training	No training		Training
	Support	No support		Support
	Active languages	Two languages		More than two languages
	Customization	No customization		Customization
	Professional environment	In-house		Freelance

Table 13 Overview of contextual inquiry of P8

### Participant P9

P9 was translating a short legal Word document from English to Dutch during the contextual inquiry. P9 uses SDL Trados to translate documents.

Participant P9 uses many keyboard shortcuts to navigate through the editor while translating. P9 uses both paper and the computer to complete the translation task. P9 knows the TEnT he is using very well, but does not customize the system because he is satisfied with the defaults. Although he might change some settings, should he know more about the TEnT. He has received training on the software and often teaches what he has learned to his colleagues. He also attends information sessions regarding the usage of the software. He is able to receive support from the developers and other colleagues.

Education	Legal background
Domain experience	Government/Constitution/Law/Politics
PC setup	Single screen, desktop
File types	Translated Document (*.xliff), Initial file type (*.docx)

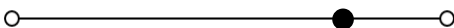
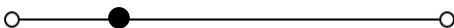

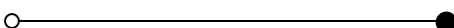




<b>Training</b>	Internal training			
<b>Workflow</b>	<ul style="list-style-type: none"> <li>• P9 receives incoming assignment on paper</li> <li>• P9 opens Translator's Desktop <ul style="list-style-type: none"> <li>◦ P9 creates local SDL Trados Studio project in TD (using template)</li> </ul> </li> <li>• P9 opens the TEnT Integration Client <ul style="list-style-type: none"> <li>◦ P9 opens project to retrieve reference documents and translation memory</li> </ul> </li> <li>• P9 opens SDL Trados Studio and opens project</li> <li>• P9 selects the document reference number <ul style="list-style-type: none"> <li>◦ P9 searches for reference number in Eur-Lex (Ctrl+Alt+O)</li> <li>◦ P9 copies the law number</li> <li>◦ P9 searches for law number in Euramis</li> <li>◦ P9 selects the source and target language <ul style="list-style-type: none"> <li>▪ Translation memory is sent via e-mail</li> <li>▪ When e-mail arrives, P9 imports the translation memory in SDL Trados Studio</li> </ul> </li> </ul> </li> <li>• P9 runs through each section and translates <ul style="list-style-type: none"> <li>◦ If the translation provided by fuzzy match or machine translation is not correct <ul style="list-style-type: none"> <li>▪ P9 refers to Eur-Lex to see previously translated elements</li> <li>▪ P9 refers to legal document to find correct terminology</li> <li>▪ P9 searches Quest Metasearch (Ctrl+Alt+Q)</li> <li>▪ P9 searches in translation memory to find similar translated words (F3)</li> </ul> </li> <li>◦ If the translation is correct <ul style="list-style-type: none"> <li>▪ P9 confirms the translation (Ctrl+Alt+Enter)</li> </ul> </li> </ul> </li> <li>• If the document has been translated <ul style="list-style-type: none"> <li>◦ P9 syncs memory and result with server using the TEnT Integration Client</li> <li>◦ P9 notifies the secretariat of the completed assignment</li> <li>◦ Colleagues of P9 review the translation and provide feedback on paper (No final sync of translation memory!!)</li> </ul> </li> </ul>			
<b>Profile</b>	Interaction technique	Mouse		Keyboard
	Media	Paper		Digital
	Software knowledge	Limited		Extensive
	Training	No training		Training
	Support	No support		Support
	Active languages	Two languages		More than two languages
	Customization	No customization		Customization
	Professional environment	In-house		Freelance

Table 14 Overview of contextual inquiry of P9



## Observational Conclusions

The contextual inquiries (or the observations) with different language professionals helped us identify some work practices and problems related to the usage of translation technology. We structured problems and work practices based on the categories: file transfer, communication medium, software knowledge, useful existing features, malfunctioning features, missing features and current problems with TEnT.

### *File transfer*

Translators use different ways to retrieve the files to complete the translation jobs. P1 and P2 receive files through email, while P3 and P4 use FTP to download the required files. P6 and P7 use a shared drive that allows not only the sharing of assignment files, but also the sharing of machine translation memory. P9 uses an in-house developed TEnT Integration Client that sets up the necessary environment using templates concerning settings related to the project such as location of the translation memories. FTP (P3 and P4) or similar sharing methods (P4, P6, P7 and P9) are preferred when files are too big to send through email. The file transfer method depends on the translation agency they are working with.

P5 uses a web portal developed by his client that serves as a communication channel between the translators and the client. The web portal provides an overview of the translation jobs assigned to each translator. Similarly, P9 uses a custom web portal to download his translation assignments. He combines this with the previously mentioned TEnT Integration client and retrieves machine translation memory from a different web portal specifically used in his domain.

When an assignment is finished, the participants use the same tool to send in their results to the client.

### *Communication medium*

All participants prefer to communicate by e-mail. Besides e-mail, P3 and P4 use an online Google Docs to ask terminology-related questions to their clients. The problem here is that it usually takes days to receive an answer. P6 has to communicate with people working in Japan, but due to the time difference and the long working hours, he does not have to wait long for an answer. Communication between P9 and the original translator of the document is done using Excel spreadsheets where the translator indicates errors discovered in the original documents.

### *Software knowledge*

P1, P2 and P4 only know how to use the very basic features of SDL Trados Studio, i.e. how to translate in the editor, use some keyboard shortcuts and add a translation memory. Other features are unknown to them due to the same reasons, i.e. there is no time to

discover and learn all the features, SDL Trados Studio is too complex and it contains too many features.

P3 and P5 are using other TEnTs, respectively Déjà Vu and an in-house developed TEnT provided to P5 by the client.

P5 also has some experience with SDL Trados Studio, but he hardly uses it due to the numerous features. Currently, he is translating with a custom offline editor provided by his client. This tool is less complex which makes it easier to use.

P3, P5, P6 and P9 are advanced users of their software. They also know how to fine-tune the software and change its settings to adapt it to their needs. Most of them also train and support their colleagues when they start using the translation software.

P7 and P8 are using screen translation, namely Swift, in combination with speech recognition, Dragon Naturally Speaking.

Several participants mentioned they did not know SDL Trados Studio while they were frequent users of the software.

### *Useful existing features*

Some participants mentioned useful existing features during the observation. P1, P2, P3, P4 and P5 like the status bar at the bottom of their TEnT, namely SDL Trados Studio, Déjà Vu and the in-house developed TEnT. In general the status bar displays statistics information about the overall document and the active segments that translators are working on.

P1, P4 and P5 indicate that the ‘AutoSuggest’ feature is very useful because it reduces the typing effort by showing suggestions when the user starts typing a word. For more information about the AutoSuggest feature, we refer you to section *Modules and Features, Translation Editor, Sub-segment matching* of the Study of Translation Systems.

P3 also pointed out that the lexicon generator and the termbase are very useful features in Déjà Vu.

P5 found the following features practical: add notes, add bookmarks, show a preview of final document in PDF version, define a certain timespan to auto save translation, define which matches need to be accepted or rejected and to free memory using a shortcut.

### *Less appreciated features*

P2 and P4 mentioned that SDL Trados Studio often provides too many tags in segments which makes it difficult to read and translate the text. SDL Trados Studio uses tags to represent information about the formatting and structure of the document. P2 disables the tags, which causes a loss of layout in the target text. P2 seemed not to be aware of the

fact that SDL Trados Studio offers the option to hide the tags completely without affecting the text formatting.

P1 mentioned that the ‘matching’ feature is not always working fine, e.g. if a match in the translation memory contains a capital letter but the word(s) in the source text does not have a capital letter, it shows however the match in the target text with a capital letter.

P3 disabled the ‘AutoWrite’ (similar to the aforementioned ‘AutoSuggest’) feature in Déjà Vu, because it causes problems (software performs unwanted actions) when the user clicks in the window while the software is proposing suggestions. P3 also complained about the import and export of large translation memories because the import and export takes several days.

P6 mentioned the in-house developed QA tool detects a large amount of false positives. This was clearly visible during the observation as the participant spent more time ignoring false positives than correcting wrongly translated words. To cope with this issue, the user explicitly has to add regular expressions to the profile to ignore a certain type of false positives.

### *Missing features*

P1, P2 and P3 expressed their wish of having the WYISYG (Live Preview) feature incorporated into their tools. This feature would enable them to detect any layout or formatting problems in the target document while translating.

It seems that a basic spelling and grammar checker has not yet been integrated in some translation environments. This is the case of the software used by P4 and P5 who prefer spell-checking their final target document in MS Word.

P4 also named some features that she was missing in SDL Trados Studio, i.e. a feature to search for a translation based on an image, a feature that proposes suggestions for acronyms, a feature that provides a link to terminology websites by a single click, a feature that provides an online chat window in the TEnT software to use while working on a project together with other translators, a feature that provides context about terms or acronym.

P5 would like to have the possibility to create a termbase and perform concordance searches in the client’s translation memory databases directly from his offline editor interface. At the moment, this is possible only via the client’s web portal.

P7 and P8 use software for captioning videos. This software does not provide any features that TEnTs offer, such as spell check, machine translations and language profiles. A great deal of manual labour has to be done in order to complete the assignment.

### *Current problems with TEnTs*

P1, P2 and P4 both mentioned that the error messages in SDL Trados Studio are unclear.

We noticed during the observation that P1 prints the PowerPoint presentations on paper to check the layout of each slide while translating. P1 and P2 also prefer reviewing the final translation on paper because it reads better. This forces them to make these changes twice, once on paper and once in the TEnT.

P1, P4 and P5 switch regularly between different applications, i.e. the source document, the TEnT, online term bases or online dictionaries, on one screen while they are translating. P4 has two screens, but one (laptop) screen just shows a duplicate of the content on the other (desktop) screen.

P5 mentioned a problem with interchangeability of documents between TEnTs. The software he is working with provides its own file type which will not be recognized as a valid file type in other TEnT.

P1, P2 pointed out that SDL Trados Studio makes small layout changes in the target text compared to the source text.

P6 uses in-house developed QA tool which is becoming deprecated. This software has not been extended or debugged in quite some time. This leads to issues with newer operating systems.

## Personas

Based on the workflows and profiles that were defined in the previous sections, personas (Pruitt & Adlin, 2005) were created. A persona is a hypothetical character which represents end-users. Our personas represent users of TEnT. We distinguish two different types of TEnT users, which are considered in the two personas we defined, each of them has specific characteristics related to the work practices during a translation job.

### *Persona 1*

Kim is 43 years old and she is working 20 years as an in-house translator. Kim's native language is Dutch and she obtained a master's degree in Germanic languages. Currently, she translates mainly from English to Dutch, but very rarely, translation jobs require a translation from French or German to Dutch. Kim usually translates business and marketing related documents.



Five years ago, she started working with TEnT, because the company she is working for wanted to improve the translation process. She followed a one day training organized by the company to learn how to work with SDL Trados. The training was given by a SDL Trados trainer. The company decided to use SDL Trados, hence Kim had to use the same tool. She made significant progress, currently translating 10 pages a day on average, while in

the beginning she was able to translate 5 pages a day on average without using a TEnT.

Kim receives a translation request through e-mail. She downloads the source document (MS Office document) and saves it on her local drive. She starts SDL Trados and begins translating the document she received. Kim does not need to import a translation memory because she is always using the same one. She consults the in-house developed term base when she is not sure about a certain translation. After each translated segment, she confirms the translation by using the default keyboard shortcut. Kim has not changed the default settings because she does not know how to do that.

After a translation is finished, Kim usually prints the translation on paper. She reads the text and corrects the mistakes on paper. Sometimes her colleague also reads the translation and marks the mistakes on paper. Kim updates the translation in the TEnT by correcting the mistakes that were marked on paper. The final translation is sent back to the requester. Finally she also moves the e-mail she got with the translation request to a folder entitled 'finished translations'.

To keep track of her translation projects, Kim enters the translation job details in a spreadsheet.

### *Persona 2*



Geert is 50 years old and is very passionate about linguistics, translation technology and terminology. He is working as a freelance translator for almost 25 years, but he also has some experience as an in-house translator. Geert always works from his home office. When he has a very strict deadline, he also spends time in the weekend to work on the translation project. Geert has experience with different TEnTs, but he prefers to use Déjà Vu. The client he is usually working for, forces him to use SDL Trados. His knowledge of TEnTs in general is very extensive. He customized the default settings of both TEnTs to adapt them to his own needs, i.e. colours of different match types, keyboard shortcuts and fuzzy match percentages.

Geert usually accepts the translation job through email and he gets the documents related to the translation through an online portal that is provided by the client he is working for. Geert works in a similar way for every translation job. He first creates some specific folders (folder with source documents, folder with invoice details, folder with untranslated documents, folder with translated documents ...) on his local drive. He calculates in a spreadsheet which files require the most work (word count) before he starts with the translation and he creates a planning. He always starts with the file that is most likely to require most of the work. He performs the pre-translation in an outdated

version of SDL Trados, because his client forces him to use it. He imports the pre-translated file afterwards in Déjà Vu to start the actual translation. Geert uploads the final translation to the online portal when the complete project is finished.

Besides translation, Geert is passionate about terminology and he has been trying to define his ideal term base fields.

He frequently attends conferences to keep up to date with the latest translation technology developments.

## Conclusions

The nine contextual inquiries revealed current work practices and issues related to the translation process of professional translators when working with a TEnT.

We detected during the observations a lack of knowledge by nearly half of the observed participants. TEnTs are seen by most of the participants as a complex tool to work with. Some explanations can be found in other observations: All observed participants had no or very limited training to learn how to work with the tool which can relate to the fact that their knowledge is limited. Most of them mentioned that they have no time to figure out how the TEnT works because they mainly work with very strict deadlines.

Profiles of the participants who have a quite similar workflow, showed a lot similarities too.

We want to highlight some additional observations about the workflow:

- Some translators make use of paper documents during the translation process, both as input and as output for the translation process, i.e. to have a physical copy of the text and layout of the source file that needs to be translated and also to read and correct mistakes afterwards when the translation is finished. The consequence of this way to make corrections is that corrections need to be made twice (on paper and in TEnT).
- We can also state that the translators use several different resources besides the TEnT to find the correct translation. These resources can be digital (such as websites) or physical (such as dictionaries).
- More than half of the observed participants have a single screen where the others are using a dual screen. The single screen set up could be improved because it was remarkable during the observation that they switch very often between different windows on one screen.
- All the participants use keyboard shortcuts provided by the TEnT to perform some actions while translating. Some translators use more shortcuts than others but all of them confirm translated segments by a certain keyboard shortcut. Almost all translators use the default provided keyboard shortcuts.

Based on our analysis of the contextual inquiries, we defined two personas. One persona is an in-house translator, less experienced with technology and using only few features of her translation environment tool. In addition, this persona has limited access to in-house TEnT training. In contrast, the second persona is a freelance translator, not only very passionate

about linguistics but also knowledgeable about translation technology and terminology management, always seeking a way to improve his translation process.

## Final Conclusions

We can conclude that the added-value of this research lies in the fact that it looks at the translation process from different perspectives: Translation Theory and Practice, LSP Perspective, translators' perspective and the engineering perspective.

The Literature study in the first part of the report aims at placing current research within the field of Translation Studies: WP5 can definitely contribute to previous research on the topic of ergonomics and usability of translation environments.

The feature inventory shows that TEnTs can be either very complex or very simple depending on the developers' vision behind the tools. However, we can distinguish a core set of modules: project management, text editor, translation memory and terminology management, integration with MTs and other external resources, QA tools. With regard to features, each developer may adopt a different approach when implementing a common feature: e.g. word counts, sub-segment matching.

The results from both the web survey and the field research are in line with what previous empirical research and case studies have shown: TEnTs are used to ensure terminology consistency, save time, increase productivity and quality.

With regard to TEnTs usage, some are more preferred than others depending on the user's profile, experience, professional needs and the user-friendliness of the tool. Moreover, due to the tight deadlines, users do not have time to dive into the different modules and features of the TEnTs to discover their real potential.

Most important areas within a TEnT that need improvements are: interoperability between various TEnTs, terminology management, language support, TM/MT integration, translation editing features, the live preview function, the user-friendliness of the UI, project management and customization options.

Further, the interviews show that there is still reluctance to embrace the cloud-based TEnTs due to privacy concerns. Moreover, the collaborative translation projects are avoided as much as possible in order to minimize the risk of inconsistencies during the translation process. On one hand, cloud-based translation management platforms are seen as an opportunity to improve the customer-vendor relationship through their portal functionality, and on the other hand, there is fear that too much automation may affect this relationship.

The interviews and contextual inquiries also provided additional insight on the work practices and needs of translators, including existing variations and commonalities, which can be exploited in further research with respect to personalisation of the translation environment and novel information visualisations within these environments.



The results of both studies will be further analysed and consolidated during the project and they may be useful not only for the creation of personalised interfaces for translation work, but also to improve the existing teaching and evaluation methodologies of translation environments.

## Acknowledgements

We would like to thank all the organizations and language professionals that participated in our web survey, semi-structured interviews and contextual inquiries. Their contribution has proved to be invaluable for this first deliverable within the WP5 work package.

## APPENDIX A: Summary of previous empirical research on TEnTs

Author	Research topic	Research method	Major findings
<b>Technology-oriented</b>			
Colominas (2008)	The usefulness of sub-sentential segmentation	Usability study of two TM systems: technical test runs of recall and precision	Sub-sentential segmentation shows better recall
Ribas (2008)	The effect of the errors contained in the TM on the final translation quality	Post-editing experiments with the students and translation professionals	If there are errors contained in the TM database, they will propagate other errors. TM may increase translators' productivity and it is useful for terminology search
<b>Workflow-oriented</b>			
Lange/Bennett (2000)	The effect of the TM/MT integration on the translation process	Case study	Translators' productivity may increase with 50–60% when they use both TM and MT provided that translators have a positive attitude towards MT.
<b>Translation-theoretical</b>			
Christensen (2003)	The usefulness of TM for legal translation	Technical runs with corpora analysis	Combine TM with a reference corpus and an authoring memory to facilitate better matches when translating legal type of text.
Dragsted (2004, 2006)	The relation between sentence-based segmentation and translators' cognitive processes.	Key-stroke logging (Translog) and eye-tracking (Tobii) plus verbalisations	TM systems need to be adjusted to match translators' natural segmentation
Fulford/Graneli-Zafra (2005)	The adoption of UK freelance translators of TM systems	Large-scale questionnaire survey	At that time, only 28% of the respondents were familiar with TM systems
Dillon/Fraser (2006)	Translators' attitudes towards TM technology	Large-scale questionnaire survey	The level of IT competence combined with a lack of awareness of the benefits of the TM might influence translators' adoption of TM
Lagoudaki (2006)	Translators' attitudes towards TM technology	Large-scale questionnaire survey	<ul style="list-style-type: none"> <li>- More than 82% of the respondents were using TM!</li> <li>- Translators use TM mainly because it saves time, it insures terminology consistency and it improves quality.</li> <li>- 16% of the translators owned a TM but they did not use it.</li> <li>- Most used tools were Trados, Wordfast, Déjà vu, SDLX and Star Transit.</li> <li>- Déjà Vu was the most user-</li> </ul>

			friendly tool, followed by Wordfast and Trados.
Wallis (2006)			- Translators seemed to prefer working in interactive translation mode rather than in the pre-translation mode.
O'Brien (2006, 2006a, 2008)	Translators' cognitive effort during post-editing TM and MT matches	Eye-tracking, screen-capture and verbalisations	<ul style="list-style-type: none"> <li>- Fuzzy TM matches increase the cognitive effort;</li> <li>- Repairing an MT match has the same impact on the cognitive effort as a 80-90% fuzzy match.</li> </ul>
Guerberof (2008, 2012)	The impact of the TM/MT integration in the translation process with focus on productivity and final quality	PE experiments with professional translators	<ul style="list-style-type: none"> <li>- Translators may reach a higher productivity level and better quality then when using TM or when translating with no aids</li> <li>- Translators level of experience may influence the PE productivity and quality</li> </ul>

## APPENDIX B: Glossary of terms

### desktop-publishing

The creation of documents using page layout skills on a personal computer. Desktop publishing software can generate layouts and produce typographic quality text and images comparable to traditional typography and printing. (Source:Wikipedia)

### globalization (G11n):

The process of making all the necessary technical, financial, managerial, personnel, marketing, and other enterprise decisions necessary to facilitate localization. (LISA, 2003)

### internationalization (I18n)

The process of generalizing a software product (or website) so that it can handle multiple languages and cultural conventions without the need for redesign. Internationalization takes place at the beginning of the program design and documentation development process. (LISA, 2003)

### localization

The process of adapting a product (software, websites) to a specific locale, i.e. to the language, cultural norms, standards, laws and requirements of the target country. (Source: eCoLoRe Glossary)

### machine translation (MT)

A translation productivity tool that works by breaking down sentences or other text segments, analysing them in context and then recreating their meaning in the target language. Machine translation works best on large volumes of well-written texts from narrow subject areas. (LISA, 2003). Bowker (2002) emphasizes that a human may assist in the process through such tasks as pre-or post-editing, but it is the computer, rather than the human, that produces an actual draft translation.

### OSCAR

A LISA Special Interest Group, which stands for “Open Standards for Container/Content Allowing Re-use”, comprises a group of localization clients, service providers and academics. The group was responsible for LISA’s TMX (Translation Memory Exchange) standard released in 1998.

### post-editing (PE)

The examination and corrections of the text resulting from an automatic or semi-automatic machine system (machine translation, translation memory) to ensure it complies with the natural laws of grammar, punctuation, spelling and meaning. (Draft of European Standard for Translation Services prEN15038: Brussels, 2004). (Source: trans/k)

**segment**

A predefined unit of a text that can be aligned with its corresponding translation. (Source: Bowker, 2002).

**sub-segment match**

A match between two chunks of text in a translation memory; these chunks do not constitute a segment or a term. (Bowker, 2002)

**translation memory (TM)**

Computer-aided translation program that stores translated sentences (translation units or segments) with their respective source segments in a database (usually called the “memory”). For each new segment to be translated, the program scans the database for a previous source segment that matches the new segment exactly or approximately (fuzzy match) and, if found, suggest the corresponding target segment as a possible translation. The translator can then accept, modify or reject the suggested translation. (Source: eColoRe)

**TMX**

Standardized translation memory exchange format, designed (under the auspices of LISA) to allow easier exchange of translation memory data between tools and/or translation vendors with little or no loss of critical data during the process. Supported by the latest versions of most leading translation memory programs. (Source: eColoRe, LISA)

**terminology extraction (TE)**

The creation of a corpus of monolingual or multilingual subject-specific terminology by extracting individual terms and phrases from a body of text. (Source: eColoRe)

**term extraction tool**

A computer program that provides functions to assist with or automate the extraction of terminology from a body of text. (Source: eColoRe)

**translation Memory eXchange (TMX)**

A standard data-exchange format for translation memories that was developed by LISA's OSCAR. The standard designates how segments of text are defined and aligned within translation memories, thus allowing these segments to be ported to any other systems supporting the standard. The TMX standards can be downloaded from <http://www.gala-global.org/lisa-oscar-standards>.

### **translation Management System (TMS)**

Sometimes also known as a globalization management system, a TMS automates localization workflow to reduce the time and money employed by manpower. It typically includes process management technology to automate the flow of work, and linguistic technology to aid the translator. (Source: Multilingual Magazine – Glossary<sup>55</sup>)

### **translation Agency**

An agency that provides translation services and manages translation projects for clients. Some translation agencies may also provide interpreters, multilingual desktop publishing (DTP) and other language related services such as website translation and software translation.

Synonyms: *Localization Agency, Language Service Provider, Translation Service Provide, Globalization Services Provider*<sup>56</sup>

---

<sup>55</sup> <http://www.multilingual.com/glossary.php>

<sup>56</sup> Source: <http://blog.globalizationpartners.com/difference-between-translation-agency-localization-services-and-lsp.aspx>

## APPENDIX C: Web Survey

The following list the questions used in the web survey. Each item displays the question as they were presented to the respondents. Bullets identify the options a respondent was able to choose to answer the question. If a question does not list any bullets or a bullet contains the “other” option, respondents were able to answer free text. All questions except 3, 6, 9, 10, 11 and 19 were mandatory.

Some questions were only presented to respondents conditionally. Questions 9, 10 and 11 were only presented to respondents that answered either “*Freelance translator*”, “*In-house translator*”, “*Post-editor*” or “*Terminologist*” to question 5 related to job description. Question 12 was only presented to respondents that answered “*In-house translator*”, “*Post-editor*” or “*Terminologist*” to question 5 related to job description. Question 13 was only presented to respondents that answered “*In-house translator*”, “*Reviser/editor*”, “*Post-editor*”, “*Terminologist*” or “*Project manager*” to question 5 related to job description. Questions 15 to 21 were only presented to respondents that answered “*Yes*” to question 14 related to the usage of CAT tools.

1. What is your age?
  - 18–24
  - 25–34
  - 35–44
  - 45–54
  - 55–64
  - 65 or older
2. What is your gender?
  - Female
  - Male
3. What country do you live in?
4. What educational program have you completed?
  - Professional bachelor in Translation
  - Master in Translation
  - Master in Interpreting
  - Master in languages (or equivalent)
  - Other
5. Select the profession(s) that best describes your job:
  - Freelance translator
  - In-house translator
  - Interpreter
  - Reviser/editor
  - Post-editor

- Terminologist
  - Project manager
  - Other
6. What language pairs do you work with?
7. How many years of work experience do you have in your field of expertise?
- 0–2 years
  - 3–5 years
  - 6–10 years
  - 10–20 years
  - 20+ years
8. Please indicate your field(s) of expertise?
- Business, Commerce, Management and Marketing
  - Accounting/Finance/Banking/Economics
  - IT, Data processing, Computing, Electronics, Control Engineering
  - Telecommunications
  - Motor vehicles, Road/Rail transport
  - Medicine and Pharmaceuticals
  - Other Mechanical Engineering
  - Insurance
  - Agriculture, Fisheries, Food, Biology, Veterinary Science
  - Tourism and Leisure
  - Civil Engineering and Construction, Mining
  - Environment/Environmental engineering and sciences/Water
  - Government/Constitution/Law/Politics
  - Education and Training
  - Electrical Engineering (including power generation)
  - Chemistry, Chemical Engineering, Nuclear Engineering, Materials Science
  - Oil and Gas Industry
  - Textiles and Clothing
  - Arts and Performing Arts
  - Shipping and Maritime Affairs
  - Aviation and Aerospace
  - History/Humanities
  - Domestic arts
  - Other
9. Please indicate the number of translation assignments that you complete on average per week?
- 1–2 assignments
  - 3–5 assignments
  - 6–10 assignments



- 11–15 assignments
  - 15+ assignments
10. Please indicate the number of pages you translate on average per assignment?
- 1–5 pages
  - 6–15 pages
  - 16–30 pages
  - 30+ pages
11. How many words on average per day do you translate?
12. Do you usually translate assignments individually or in a group?
- Individually
  - In group
  - Both
13. Do you ever work remotely (not directly from your office)?
- Yes, (almost) all the time
  - Yes, sometimes
  - No
14. Do you use a CAT tool?
- Yes
  - No
15. How long have you been working with CAT tools?
- 0–2 years
  - 3–5 years
  - 6–10 years
  - 10+ years
16. What CAT tool(s) are you currently using?
- SDL Trados
  - memoQ
  - Déjà vu
  - Wordfast
  - OmegaT
  - CaféTran
  - Star Transit
  - XTM
  - Memsource
  - Wordbee
  - Across
  - Other
17. How often do you use the CAT tool you marked in the previous question?
- Occasionally
  - Part-time

- Full-time
- What CAT features do you use on a regular basis?
- Project management features
- Processing of complex file formats
- Translation memory
- Pre-translation with TM
- Pre-translation with MT
- Terminology management
- Term extraction feature
- Concordance feature
- Alignment feature
- Analysis/Statistics/Word counts
- QA features
- Review features
- Customization features
- Collaborative features (TM sharing, instant chat, etc.)
- Other

18. Why do you prefer working with CAT tools?

- Quality improvement
- Time efficiency
- Productivity increase
- Terminology consistency
- Ease of collaboration
- Other

19. What features need to be improved or added in your current CAT tool to optimize your translation process?

20. How did you learn to use CAT tools?

- Video tutorials
- (In-house) training
- Workshops
- Webinars
- During my education
- Self-taught
- Other

21. Which aspects are important for you when using a CAT tool?

- Easy to learn
- Easy to use
- Easy to customize
- Speed
- Resource management (translation memories, terminology, reference files)

- Quality Assurance checks
- Project management
- Compatibility with other tools
- Other

## References

- Alabau V., J.González-Rubio, L.A.Leiva, D.Ortiz-Martínez, G.Sanchis-Trilles, F.Casacuberta, B.Meso-Lao, R.Bonk. M.Carl, & M.García-Martínez (2013): User evaluation of advanced interaction features for a computer-assisted translation workbench. *Proceedings of the XIV Machine Translation Summit*, Nice, September 2–6, 2013; ed. K.Sima'an, M.L.Forcada, D.Grasmick, H.Depraetere, A.Way; pp.361–368.
- Asare, E.K. (2011). An ethnography study of translation tools in a translation agency: implications for translation tool design.
- Austermühl, Frank. (2001). *Electronic Tools for Translators*. Manchester: St. Jerome Publishing, 2001.
- Beyer, H., & Holtzblatt, K. (1998). *Contextual Design: defining customer-centered systems*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.
- Le Blanc (2013). Translators on translation memory (TM). Results of an ethnographic study in three translation services and agencies. *T& I-The International Journal of Translation and Interpreting Research*. Vol 5, No 2. URL: <http://www.trans-int.org/index.php/transint/article/view/228>. Last accessed in January 2015.
- Bowker, Lynn (2002). *Computer-Aided Translation Technology*. Ottawa: University of Ottawa Press.
- Bowker, Lynn (2012). *Managing terminology for translation using translation environment tools: towards a definition of best practices*. Ottawa: University of Ottawa Press.
- Beyer, H., & Holtzblatt, K. (1998). *Contextual Design: defining customer-centered systems*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.
- Christensen, T. P., & Schjoldager, A. (2010). Translation-memory (TM) research: What do we know and how do we know it? *Hermes – The Journal of Language and Communication*, 44, 1–13.
- Christensen, T. P. (2011). Studies on the Mental Processes in Translation Memory-assisted Translation – the State of the Art 2 TM-assisted Translation as a Mental Activity, 137–160.
- Colominas, C. (2008). Towards chunk-based translation memories. In *Babel* 54(4), 343–354.
- Désilets, A., Brunette, L., Melançon, C., & Patenaude, G. (2008). *Reliable Innovation : A Tecchie's Travels in the Land of Translators*. *Proceedings of the AMTA 2008*. Waikiki, Hawaii, October 21–25, 2008.

Désilets, Alain; Patenaude, Genevieve; Sirois, A. (2009). Tutorial : Up close and personal with a Translator – How Translators Really Work.

Dillon, S., & Fraser, J. (2006). Translators and TM: An investigation of translators' perceptions of translation memory adoption. *Machine Translation*, 20(2), 67–79. doi 10.1007/s10590-006-9004-8

Dragsted, B. (2006). Computer-aided translation as a distributed cognitive task. *Pragmatics & Cognition*, 14(2), 443–464.

Dragsted, B. (2004). Segmentation in Translation and Translation Memory Systems – An Empirical Investigation of Cognitive Segmentation and Effects of Integrating a TM System into the Translation Process. Copenhagen: Samfundslitteratur.

eCoLoRe. (2012). Glossary of Terms Related to eContent Localisation. University of Leeds, Centre for Translation Studies. <http://ecolore.leeds.ac.uk/xml/materials/overview/glossary.xml?lang=en>

eCoLoRe. (2012). eContent Localisation Resources for Translator Training. <http://ecolore.leeds.ac.uk/>. [Last Accessed January 2015]

EAGLES Evaluation Working Group (1999) 'The EAGLES 7-step recipe' online available: <http://www.issco.unige.ch/en/research/projects/eagles/ewg99/7steps.html> [accessed 2 Nov 2014] EC (2009) report on the size of the language industry in the EU.

Ehrensberger-Dow, Maureen & Gary Massey. (2014). Translators and machines: working together. In: *Man vs. Machine? Vol. I. Proceedings of XXth World Congress of the International Federation of Translators*, Berlin, 4–6 August 2014, 199–207

Feder, M. (2001). Computer Assisted Translation. A Proposal for Tool Evaluation Methodology. School of English, Adam Mickiewicz University, an unpublished Ph.D. thesis. Retrieved on January 15, 2015 from <http://www.transsoft.seo.pl/>.

Fulford, H./Granell-Zafra, J. (2005). Translation and Technology: a Study of UK Freelance Translators. In *Journal of Specialized Translation* 4, 2–17.

Guerberof, A. (2008). Post-editing MT and TM. A Spanish case. In *Multilingual Computing* 19(6), 45–50.

Gow, F. (2003) Metrics for Evaluating Translation Memory Software (M.A.), University of Ottawa.

Garcia, I. (2007). Power shifts in web-based translation memory. *Machine Translation*, 21, 55–68. doi 10.1007/s10590-008-9033-6

García, I. (2009) 'Research on translation tools', *Translation Research Projects* 2, 27.

Höge, M. (2002) Towards a Framework for the Evaluation of Translators' Aids Systems (PhD), Faculty of Arts, Department of Translation Studies, University of Helsinki.

ISO/IEC 25010: Systems and software engineering -- Systems and software Quality Requirements and Evaluation (SQuaRE) -- System and software quality models (2011).

Lange, C.A./Bennett, W. eld Scott. (2000). Combining machine translation with translation memory at Baan. In *Sprung, Robert C. (ed.) Translating into success*, Amsterdam/Philadelphia: John Benjamins, 203-218.

Lazar, J., Feng, J. H., & Hochheiser, H. (2010). Research Methods in Human-Computer Interaction. Wiley.

Likert, R. (1932). A technique for the measurement of attitudes. *Archives of Psychology*, 22, 1-55.

Lagoudaki, E. (2006). Translation Memories Survey 2006. User's perceptions around TM use. *Translation and the Computer*, 28, London, England: ASLIB, 1-29. Retrieved from <http://mt-archive.info/Aslib-2006-Lagoudaki.pdf>

LISA (2003). The Localization Industry Primer, second edition. Revised version by Arle Lommel. Retrieved from <http://www.dynamiclanguage.com/home14/wp-content/uploads/2014/07/LISA-Globalization-Primer.pdf> on March 2015.

Lossner, Kevin (2013). The 2013 translation environment tools survey. <http://www.translationtribulations.com/2014/01/the-2013-translation-environment-tools.html>. Last accessed, March 2015.

Karen Holtzblatt, Jessamyn B. Wendell, and Shelly Wood. (2004). Rapid Contextual Design: A How-to Guide to Key Techniques for User-Centered Design (Interactive Technologies). Morgan Kaufmann Publishers Inc.

Makoushina, Julia (2007). Translation Quality Assurance Tools: Current State and Future Approaches. Retrieved on January 2015 from <http://mt-archive.info/Aslib-2007-Makoushina.pdf>.

Massey, G. and Ehrensberger-Dow (2011). Technical and Instrumental Competence in the Translator's Workplace: Using Process Research to Identify Educational and Ergonomic Needs." ILCEA [online] available on June 30<sup>th</sup>, 2011. Last accessed January 2015. URL: <http://ilcea.revues.org/1060>

Melby, A.K.; Smith M.R.; Peterson, J. (1980). "ITS: interactive translation system", *Coling 80: Proceedings of the Eighth International Conference on Computational Linguistics*, Tokyo, Japan, pp. 424-429. Retrieved on December 19, 2014 from <http://acl.ldc.upenn.edu/C/C80/C80-1064.pdf>.

- Moorkens & Sharon O'Brien (2013). User attitudes to the post-editing interface. *Proceedings of MT Summit XIV Workshop on Post-editing Technology and Practice (WPTP-2)*, Nice, France, 2, September 2013; pp. 19–25.
- Mossop, B. (2006). Has computerization changed translation? *Meta*, 51(4), 787–793. doi 10.7202/014342ar
- Quah, C. K. (2006). *Translation and Technology*. Basingstoke, UK: Palgrave Macmillan.
- Robinson,
- Vieira & Specia (2011). A review of translation tools from a post-editing perspective. *Proceedings of the Third Joint EM+/CNGL Workshop "Bringing MT to the User: Research Meets Translators" (JEC '11)*, Luxembourg, 14 October 2011; pp.33–42. [PDF, 801KB]
- Nogueira, D. (2004). In Memoriam: Emilio Benito 1947—2004. In *Translation Journal*. 8–2 .Retrieved from: <http://translationjournal.net/journal/28benito.htm> [Last accessed 2 January 2015]
- O'Brien, S. (2007). Eye-tracking and translation memory matches. *Perspectives: Studies in Translatology*, 14 (3), 185–205. doi 10.1080/09076760708669037
- O' Brien, S. (2008). Processing fuzzy matches in translation memory tools: An eye-tracking analysis. *Copenhagen Studies In Language*, 36, 79–102.
- O'Brien, S., Hagan, M. O., & Flanagan, M. (2010). Keeping an eye on the UI design of Translation Memory : How do translators use the “Concordance” feature ? *28th Annual Conference of the European Association of Cognitive Ergonomics*, 187–190.
- Plitt, M. & Masselot, F. (2010). A productivity Test of Statistical Machine Translation Post-Editing in a Typical Localisation Context. *The Prague Bulletin of Mathematical Linguistics No. 93.2010*. pp.7–16. ISBN 978–80–904175–4–0. doi:10.2478/v1018–010–0010–x.
- Pym, A. (2011). What technology does to translating. *The International Journal for Translation and Interpreting Research*, 3(1), 1–9. Retrieved from <http://www.trans-int.org/index.php/transint/article/view/121/81>
- Pym, A., & Biau Gil, J. R. (2006). Technology and translation (pedagogical overview). In A. Pym, A. Perekrestenko, & B. Starink (Eds.), *Translation Technology and its Teaching* (n.p.), Tarragona, Spain: Universitat Rovira I Virgili and Intercultural Studies Group.
- Anthony Pym (2011). Translation research terms: a tentative glossary for moments of perplexity and dispute. In *Ed. Translation Research Projects 3. Tarragona: Intercultural Studies Group*, 2011. Retrieved, January 11, 2015 from [http://isg.urv.es/publicity/isg/publications/trp\\_3\\_2011/pym.pdf](http://isg.urv.es/publicity/isg/publications/trp_3_2011/pym.pdf)

Pruitt, J., & Adlin, T. (2005). *The Persona Lifecycle: Keeping People in Mind Throughout Product Design*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.

Reinke, U. (2013). State of the Art in Translation Memory Technology. *Translation: Computation, Corpora, Cognition*, 3(1). Retrieved from <http://www.t-c3.org/index.php/t-c3/article/view/25>

Rahzeb Choudhury and Brian McConnell (2013). *TAUS Translation Technology Landscape Report*. Published by TAUS BV, De Rijp.

Wassmer, T. Software reviews. Atril Déjà Vu. Retrieved from: <http://www.softreviews.org/Localize2/dejavu/dejavu.html> [Last accessed 2 January 2015]

Zerfass, A. (2002) ‘Comparing Basic Features of TM Tools’, *Multilingual Computing & Technology*, 13 (7), 11–14.

Zerfass, A. (2010) *MemoQ 4*, *Multilingual Computing & Technology*, 21 (4), 14–17.

Zetzsche (2014). *Translator Tool Box, A computer primer for translators* (version 11.0, 2014)

## Websites

Déjà Vu X 7.5. (2013). Retrieved from: <http://taalunieversum.org/taal/terminologie/tools/software.php?id=133> [Last accessed 29 December 2014]

Leaders in Computer Assisted Translation since 1993. (2013) Retrieved from: <http://www.atril.com/content/atril> [Last accessed 29 December 2014]

Kilgray Technologies: <https://www.memoq.com/>

MateCat: <https://www.matecat.com/>

OmegaT: <http://www.omegat.org/>

SDL: <http://www.sdl.com/about/sdl/history/>

Wordbee: <http://www.wordbee.com/>