UNIVERSIDAD DE EL SALVADOR
FACULTAD DE CIENCIAS Y HUMANIDADES
ESCUELA DE POSGRADOS
DEPARTAMENTO DE IDIOMAS EXTRANJEROS

Comparative Analysis between two leading Computer Assisted Translation Tools used by Student Translators at Universidad Evangélica de El Salvador

To opt for:

Master's degree in translation and interpretation of English to Spanish - Spanish to English

STUDENTS:

Emerson Mario Ovidio Sánchez Ruiz (SR00014)

Marcela Melany Melara de Sánchez (MH04064)

Advisor:

MT. Ricardo Gamero

Ciudad Universitaria, Viernes 19 de Febrero de 2016

Authorities of the University of El Salvador

José Luis Argueta Antillón (interino)
RECTOR

ACADEMIC VICE RECTOR

Carlos Armando Villalta Zavaleta (interino)
ADMINISTRATIVE VICE RECTOR

Ana Leticia Zavaleta de Amaya (interina)
GENERAL SECRETARY

AUTHORITIES OF THE SCHOOL OF ARTS AND SCIENCES

José Vicente Cuchillas Melara
DEAN

Edgar Nicolás Ayala
VICE DEAN

Rafael Ochoa Gómez
SECRETARY OF THE SCHOOL

Acknowledgements

We are grateful with The Lord for allowing us to take this step towards our professional development.

TABLE OF CONTENTS

Introduction

The need to communicate seamlessly and effectively has become significant and the translation market in El Salvador is highly competitive, demanding ever higher standards of performance and productivity both from experienced and novice translators. The requirements novice Salvadoran translators will have to meet are definitely going to increase as the demands for faster and accurate production of translated documents increase as well.

This document contains a research study aimed towards comparing two Computer Assisted Translation Software often used by student translators enrolled in the the sixth and eighth semester of the Translation and Interpretation of the English Language Bachelor´s program at Universidad Evangélica de El Salvador.

This research study is comprised of the following parts: statement of the problem, rationale, delimitation, methodology design, and design of analysis of data, conclusions, recommendations, and bibliography.

Each major component has been described in order to provide as much information as possible of the proposed relevance of each item in regards of their adaptability and other major areas.

I. Statement of the problem

The most important thing to take into account upon choosing to use a CAT (Computer Assisted Translation) tool is how it may affect the current work-flow of a translator. Such is not the case of a novice translator or student translator who is starting to develop his or her own work-flow from scratch. The wide array and variety of options to choose from range from complete suites of applications typically known as workbenches to single programs specialized in a single operation. Software prices also vary accordingly. Since most novice translators, who are in most cases just breaking into the market, are not capable to afford an expensive piece of software, to opt for a proprietary solution may not be entirely attainable due to their high price at the moment of writing this document. The Open Source movement as well as Integration of Cloud services have managed to create multiple software choices out of which two will be analyzed throughout the course of this study. This also applies to companies such as Google that render tools that provide adequate solutions to this issue for free and have also the same learning curve as Open Source Solutions. This will help us know the answer to the question, **which of these two CAT software alternatives: Google Translator Toolkit or OmegaT, used by Student Translators is more intuitive, easier to use, and better optimizes their workflow?**

II. Rationale

The increase in the volume and diversity of text types for translation has naturally led to an expansion of the skills expected of a translator and in some cases has even produced new language-related professions some of the ones mentioned by Gregory Shreve are such as:

bilingual editor, multimedia designer, research and information specialist, cultural assessor, multicultural software designer, software localizer, terminologist, or project manager. (Shreve G., 2000, p 228)

Although translation memory systems are now widely used, there is relatively little research on the impact that they have on either the way a translator works or the output they have produced using them. (Williams & Chesterman, 2002, p15.)

This research focuses on how the usage of open source translation software impacts the way translators work and how it may affect the resulting output of their endeavors in translating different types of documents as well as in a variety of formats. Most novice translators will benefit from having a thorough comparison between not only the technical specifications of each of the aforementioned CAT tools but also how real novice users perceive how each of the characteristics offered adapts better to suit their emerging needs. The practical implications of this research are:

- It depicts the results upon deploying the usage of CAT software in a group of novice translators and comparing the flexibility, responsiveness, ease of use as well as increment in productivity, if any, offered by each CAT tool.
- It establishes the main difference in terms of usability upon comparing the technical capabilities of the two CAT tools in this study.

As mentioned in previous paragraphs, very little research has been conducted on the field of how Computer Assisted Translation software affects the work-flow of seasoned translators, let alone the emerging patterns in the work-flow of novice translators. The theoretical value of this research aims towards expanding the knowledge related to the usage of these CAT tools in real world situations as well as generating a pool of suggestions in order to improve how these CAT tools operate both in their back-end (technology and engine) and front-end (UIX: User Interface Experience).

Open Source solutions in the field of translation are not new and have been in active development since the late 1990's. One, if not the most prominent characteristics of Open Source Software is not its price, which of course is free due to the restrictions of the license (either GNU, GPL) under which most of them are distributed but the way in which the community of programmers involved in their development are actively searching for ways to improve the base technologies included in their programs as well as to offer a more efficient user interface experience. This research generated several suggestions that may aid in improving both back-end and front-end characteristics of the programs mentioned before.

III. Delimitation

This research study was conducted at the School of Social Sciences at Universidad Evangélica de El Salvador during the Second semester of the academic year 2015 (between the months of July - November).

This research was conducted in a single group of novice student translators from the Fourth year of the Bachelor Degree in Translation and Interpretation at the Languages Department of Universidad Evangélica de El Salvador.

This group of Students Translators used both CAT software tools mentioned before: Google Translator Toolkit and OmegaT in order to assess which proved to be more useful, intuitive, and more efficient when it comes to operate within the emerging patterns of their workflow.

IV. Objectives

A. General Objectives:

1. To determine which of the two Computer Assisted Translation tools presented: OmegaT and Google Translator Toolkit proved to be more efficient at optimizing the workflow of Student Translators.

B. Specific Objectives

1. To Identify the technical differences that directly impact the perceived optimization in workflow by the user.
2. To generate suggestions to improve both the back-end and frontend development of these tools to create a more powerful, responsive, and flexible user experience.

V. Assumptions

The following assumptions were taken into account upon determining the methodology to use and the expected results:

- Novice Translators and Student Translators who use Computer Assisted Translation tools perceive a considerably higher increment in terms of productivity than Novice Translators and Student Translators who don't use any Computer Assisted Translation Software.
- Finalization time of a translation in the hands of a Novice Translator and Student Translator becomes increasingly affected, minimizing it, by means of using Computer Assisted Translation tools.
- Active Translation Time of a translation in the hands of a Novice Translator and Student Translator becomes increasingly affected, minimizing it, by the means of using Computer Assisted Translation tools.
- Most Computer Assisted Translation tools offer similar core characteristics such as integrated translation memory modules in order to upgrade a novice translator's productivity and other similar elements such as integrated those offered by proprietary software.

VI. Main research questions

- What are the technical back-end specifications of the following Computer Assisted Translation tools: OmegaT and Google Translator Toolkit?

- What are the front-end technical specifications (user interface and user interface experience design) of the following Computer Assisted Translation tools: OmegaT and Google Translator Toolkit?

- What are the differences regarding translation memory and terminology database management solutions between the following Computer Assisted Translation tools: OmegaT and Google Translator Toolkit?

- How does using each of the following Computer Assisted Translation tools: OmegaT and Google Translator Toolkit impact the work-flow of a Novice Translator and Student Translator in terms of:
    - Volume of work processed
    - Active Work time
    - Down Time: time in which no work is being performed due to background processes being carried out (i.e.: starting time, saving, production of final documents, preparation of documents for pre-translation and batch processes such as pseudo-translation, format converting, resources allocation and workspace processing)
    - Number of errors.

- Which of the following Computer Assisted Translation tools: OmegaT and Google Translator Toolkit is perceived by novice translators as the one that optimizes their work-flow better in terms of:

- Translation memory and Terminology Database management
- Quality assurance management: Revising /editing for accuracy
- Overall Ease of Use:
    - Intuitiveness in user interface design
    - Adaptability
    - Flexibility
    - Responsiveness
    - Inter-operability with other applications
    - Production/outcome

VII. Theoretical Framework

A. A brief history of Computer Assisted Translation Technologies

There is a difference to be made between the most common approaches to technology in translation: the first being what will be referred throughout this study as Computer Assisted Translation technologies or CAT technologies, Computer Assisted translation is the use of computers software to assist a translator in the translation process which contrasts with the second term which is Machine Translation which refers to translation being carried out mainly by computers with some intervention by humans (Bowker & Fisher, 2010). As stated by Bert Esselink:

"*Where the purpose of Machine Translation is to assume and perform many of the tasks normally completed by a translator, Computer Aided Translation Tools are used to support a translator, by eliminating repetitive work, automating terminology lookup activities, and recycling previously translated texts"* (Esselink, 2000, p. 359)

Attempting to create a fully automatized Machine Translation Software is still a daunting task but were the efforts to do so that helped create what we now know as CAT technologies. In 1964 the Automatic Language Processing Committee (ALPAC) a the US National Academy of sciences was formed and in 1966 this same committee elaborated a report that halted funding to research studies being carried out in relation to the development of automated or machine translation technologies (Palacz, 2003). Due to other instances similar to the one in 1966, researchers turned their attention to CAT technologies during the mid 1960's first with the creation of terminology banks which were basically storage of structured or paired pieces of information and later on with the arrival of modern CAT tools during the late 1970's and early 1980's which

capabilities extended to storing, retrieving and managing information (Bowker & Fisher, 2010).

The foundations for modern CAT tools were developed by visionaries such as Martin Kay (1980) but it wasn't until the 1990's when such software became distributable and commercialized as we know it today (Bowker & Fisher, 2010). During the years of 1992-1994 when major software suppliers such as IBM, GlobalWare, Eurolang, SDL international, Atril, and TRADOS, started launching their first commercial Complete Modular Software for Translators the world saw the birth of the modern Translator CAT tool workbench as we know it today. Although it is difficult to pinpoint the leading company in terms of technological prowess, many of the technologies used today in modern CAT tools can be traced back to the German company TRADOS which introduced the now ubiquitous split screen editor window and terminology management module in their first software package with the name Translator's workbench for DOS (Palacz, 2003)

There is a distinction to be made among CAT tools and other software. CAT tools differ from very specific and task oriented software such as word processors, email clients, and automated translation or terminology management software given that today they tend to integrate most of these features in what is referred as Translation Environment Tools (TEnT) or by their name given among translators: workbenches & workstations. These applications are often built around a module which is a parallel corpus construction software referred as TM or translation memory which works together with a Terminology Management System (Palacz, 2003). According to Bert Esselink:

"CAT tools can be categorized by: Translation Memory Tools, Terminology Tools, and Software Localization Tools" (Esselink, 2000, P. 360)

B. TM modules and other components of modern CAT-tools

As Lynn Bowker and Des Fisher from the University of Ottawa describe:

"A TM (Translation Memory) is a tool that allows users to store previously translated texts and easily consult them for potential reuse". (Bowker & Fisher, 2010, p.61).

A translation Memory Module works by means of segmentation *segmentation* of the source text. A segment is a text element that is considered by the application as the smallest translatable unit. A translation memory database is a repository for all these segments and their translation. (Esselink, 2000)

Translation memories have become one of the most widely used features within a translator's endeavors. There are a few ideas that need to be expanded before continuing. First is the fact that Translation Memory modules work by means of matches or parallel segments that have been paired in both their source language form and target language adaptation. Second, Translation Memory modules work by first storing these pairs and later identifying the potential match of a new segment to a previously translated pair. Therefore, Translation Memory modules can look up among a great number of segments and later categorize the match by assigning a percentage which places the pairing within the following categories:

Types of matches used in Translation Memory modules	
Exact Match	A segment in the new text is identical to a translated segment in the Translation Memory Database

Full Match - Context Match	A segment in the new text is identical to a translated segment in the Translation Memory Database except for a few context related discrepancies such as figures, names, etc.
Fuzzy Match	A segment in the new text has some degree of similarity to a translated segment in the Translation Memory Database which can range from 1% to 99%
No Match	No part of a segment in a new text is identified nor has any similarity to a segment stored in TM Database

As mentioned before, Translation Workbenches or Workstations are comprised of different modules or components TM modules being one of the plethora of features that modern CAT tools offer. The following are some of the most common components of modern CAT tools (Bower & Fisher, 2010):

Components of modern Translation Workbenches

Concordancer	Searches text for all instances of user defined character strings and displays them in context
Document Analysis Module	Compares the content of a new text to translate with those of a specified TM database or termbase to determine the number/type of matches
Machine Translation Module	Generates a machine translation of a segment not found in the TM database

Project Management Module	Helps users track client information, manage deadlines, and maintain project files for each translation project
Quality Control Module	Includes spelling, grammar, completeness, or terminology – controlled language – compliance checkers.
Term Extractor	Analyzes texts and extracts candidate terminology

C. Advantages of CAT tools in terms of productivity

Translator Workbenches offer a lot of advantages to translators being the capability of using a Translation Memory module the biggest. Translation Memory modules offer the capability of recycle previously translated text in new instances where it is possible to identify a similar segment and opt to modify the translation instead of translating from scratch. This practice alleviates human translators from retrieving information directly from their own memory and consultation is reduced to just a few instances (Palacz, 2003). In terms of quality, CAT tools still depend of human verification and proofreading. The use of CAT tools has inherently affected translators in terms of output of translated segments as well as remuneration or compensation for their endeavors. (Bowker & Fisher, 2010)

Some of the most widely perceived advantages of using a CAT tools are detailed below:

- Existing Translation Memories can be used to translate new material

- Most Workbenches have a Terminology management system that allows to retrieve key terms
- Most workbenches include key modules that offer detailed information about the document to translate such as word count and number of segments
- Several translators can work on a single project
- Training in complex word processing or coding utilities is not necessary
- According to Bert Esselink, productivity levels (on a localization project) can be increased by a 30% or even 50% and total translation costs can be reduced by 15% to 30%. (Esselink, 2000 p.365)

D. Open Source CAT tools

CAT tools are improving with time and companies that develop are including more and more sophisticated features in current releases such as linguistic analysis and the ability to recall surrounding segments of context matches (Bowker & Fisher, 2010). However one of the biggest leaps forward in the development of modern CAT tools is the current development of Open Source Translation Workbenches and Modules. Free and Open Source Software (FOSS) is gaining popularity in the translators' community due to their low operations cost and flexibility. Open Source software does not automatically mean that is free but that the source code from which the software itself originated is open and free to the public under licenses such as the GNU – GPL (General Public License) allowing individuals or organizations to modify, reuse, and incorporate the code in their own projects if so they wish. (Mckay, 2004)

The terms of the most common Open Source licenses allow individuals or organization to:

- Install the software on as many machines as he/she wants
- Allow any number of people to use the software at once
- Copy the software and give it to anyone

- Modify the software, as long as certain features are kept intact (most commonly the licensing agreement)
- Freely (in the sense of "without restrictions") and for any purpose distribute or sell the software without paying royalties to the original developer

One of the most notable examples of Open Source Software is the Linux Operating System which is actively developed and distributed under the GPL license. Open Source Software alternatives in the CAT tool market are very important to translators due to their flexibility and affordability, especially for freelance translators who often cannot afford expensive proprietary translator workbenches. However most translators are still reluctant to switch to Open Source Software due to several reasons such as: lack of knowledge about FOSS, incompatibility of existing resources, leading companies do not create FOSS compatible alternatives of their products, lack of financial support due to volunteer developing of the code, Open Source Software is generally perceived as being of less quality than their proprietary counterparts (Mckay, 2004).

D. Advantages of Free Open Source Software against Proprietary alternatives

Among the advantages perceived in using Open Source Software are: the fact that is cost-effective. In spite of the need to pay for documentation, support and training similar to proprietary software individuals are not required to pay high prices for software, upgrades, data recovery after a virus attack, tracking licensing paperwork, etc. In addition, having free or low-cost tools available for critical but unexciting tasks like backups makes it more likely that you will actually obtain the tools. Open source encourages innovation therefore allowing leading companies to improve their alternatives while volunteer developers can also take steps toward refining their offerings. Open-source licenses are less restrictive allowing users to install and redistribute the software as they please (Mckay, 2004).

One of the most notable advantages of FOSS is its affordability. A simple comparison between the leading alternatives in both markets may provide more insight into the matter. The leading proprietary software company SDL TRADOS recently released their newest workbench which includes terminology management and Translation Memory modules under the name TRADOS 2014 for a cost of over $800 for the version aimed towards freelance translation professionals whereas the FOSS alternative OMEGA T is as of today still $0 including the same features as its counterpart.

E. Leading CAT alternatives available to Student Translators

OmegaT (www.omegat.org)

OmegatT is a cost-free and open source CAT tool, currently on its 2.5 released version as of June 2014. OmegaT is written in Java, so will run on Windows, Linux, and Mac OS X. It supports the traditional CAT features such as fuzzy matching, match propagation, simultaneous use of multiple translation memories, and multiple file formats.

Among the technical specifications of this tool are the following:

- Segmentation can be configured based on language or based on file format
- Segmentation rules inherit values from each other.
- Localization is readily accessible via the following formats: Android Resource, Java properties, Mozilla DTD, Windows Resource, WiX Localization, ResX.
- Supports XLIFF format (SDL TRADOS working file extension)
- Support for TBX termbases and glossaries is readily accessible.

Google translator Toolkit (https://translate.google.com/toolkit)

Google's offer is based around the idea of allowing freelance and "non-professional" translators to access a rather simple web interface and powerful machine translation

capabilities (Garcia, 2009). Among the technical specifications of this tool are the following:

- A user interface in 36 languages: user interface is available in 36 languages, including Tamil, Bulgarian, Catalan, Croatian, Czech, Danish, Dutch, English, Filipino, Finnish, French, German, Greek, Hindi, Hungarian, Indonesian, Italian, Japanese, Korean, Latvian, Lithuanian, Norwegian, Polish, Portuguese (Brazil), Portuguese (Portugal), Romanian, Russian, Serbian, Simplified Chinese, Slovak, Slovenian, Spanish, Swedish, Thai, Traditional Chinese, Turkish, Vietnamese, Urdu.
- Its Application Program Interface, which is the core of an application is closed source.
- No localization options are available at the time of conducting this study.
- Translation Memories in TMX format are supported but not updatable at the time of conducting this study.
- Termbases are not supported at the time of conducting this study.

VIII. Methodology

This chapter details the method and the steps that allowed the previously stated research study to take place. Several aspects of the methodological framework such as the type of study, techniques used to gather and process relevant information as well as the procedures to analyze the resulting data are detailed below.

The current research study is non experimental in its nature since none of the circumstances have been manipulated and its of qualitative character since it aims to describe the effects of using specific translation software packages on the productivity of novice translators under very specific circumstances, in this case the use of CAT tools by the subjects.

Despite the intention to gather numeric or statistical data in order to assess the increment in productivity that the subjects may have over a set period of time, several non-numerical or statistical types of information will be taken into account in order to portray more distinct answers to questions such as preference, ease of use, and user friendliness of each software package.

A. Translation Software to compare

The translation software is under the GPL license and share the characteristics of being either Open Source or multiplatform.

1. OMEGAT
2. Google Translator Toolkit

B. Method and techniques

The methodology used in this study is qualitative and it's heavily influenced by the structure of comparative analysis for software evaluation which allows different characteristics that can be quantified such as the results derived from the translation endeavors of novice translators to be unified with the results from observation notes in order to have a wider notion of users' preferences and experience while using these software packages than merely just the increment in translated words per document.
There are several testing techniques used in comparative analysis for software evaluation studies. In this particular case the SCENARIO TESTING evaluation procedure which started being used in the 1990s (EAGLES, Experts, Advisory Group on Language Engineering Standards 1995:33) was chosen particularly because it aims at using a realistic user background for the evaluation of software.
As stated by Blazjec Palacz (2003):

> *"SCENARIO TESTING… It is an example of black box testing. In scenario testing the suitability of the product for everyday routines is subject to evaluation. It usually involves putting the system into its intended use by performing standard tasks. Scenario testing provides good empirical information concerning the usability of the product. It also supplies information on accuracy, adaptability, operability etc."* (Palacz, 2003, P.43)

C. Population

The population that tested and used the features of the CAT tools was comprised of 21 students translators from the 4th year of the translation and interpretation degree at Universidad Evangélica de El Salvador each assigned to operate both software packages. It is due to the fact that this study pretended to gather information related to which of the two proposed Computer Assisted Translation tools provides novice interpreters with the best features to improve their productivity and optimize their workflow that these students were deemed to be suitable due to the following characteristics. They had:

1. Basic knowledge of translation techniques
2. Limited experience on translating documents under rigorous conditions and constraints of time and precision
3. Limited contact with Proprietary code Computer Assisted Translation tools.
4. Limited experience in terms of manipulating tools designed specifically to translate documents
5. Basic knowledge of concepts and commands used upon manipulating text processing software.

D. Techniques to gather data

The data were gathered through the following means:

1. The translated files generated by the students translators

   The pieces to translate consist of 3 each one varying in nature or format (1 text documents, 1 slideshow, 1 webpage) but retaining similarities in regards to content (related to terminology management software) during a period of 6 weeks.

The documents were gathered with the purpose of analyzing the following indicators:

- Number of mistakes: in punctuation, grammatical elements, syntax.
- Correlation of format and space between the original and the translated documents
- Omissions in information (due to some tools not being able to display certain items such as notes, comments, footnotes, etc.)

2. Observation guide and notes, which should detail the following indicators to observe during the translation phase:

- Time taken to finalize each translation
- Time destined for revision, verification and proofreading of the translated documents. This is due to the fact that these options may vary or not be present in the compared tools.
- Active translation time, which is the period in which the novice translators is actively translating a document without taking into account the time taken to revise and proofread the document.

3. Survey :

A short survey was taken by each participant in order to gather information about their impressions related to the following aspects:

- How intuitive the user interface of the tool is
- Degree of difficulty experimented by participants upon using the basic functions of each tool
- Degree of difficulty experienced by users upon producing the translated document

IX. Data analysis design

The data gathered using the previous tools was analyzed by means of comparing the results retrieved from the time spent by participants translating and revising their work as well as allocating their answers in terms of frequency in order to have a general idea of which Computer Assisted Translation software package offers better features to increment the productivity of a novice translator.
In the end, the analysis of the data answered the following questions:

- Which Computer Assisted Translation software offered more adequate features to increase the productivity of a Novice or Student Translator?
- Which Computer Assisted Translation software proved to be easier to operate and more intuitive for Novice or Student Translators translators?

    - In terms of ease of navigation
    - In terms of generation of translated documents
    - In terms of project management
    - In terms of advanced features management

X. Data Analysis

A. Analysis of data related to time spent in each activity within the workflow of a Novice or Student Translator

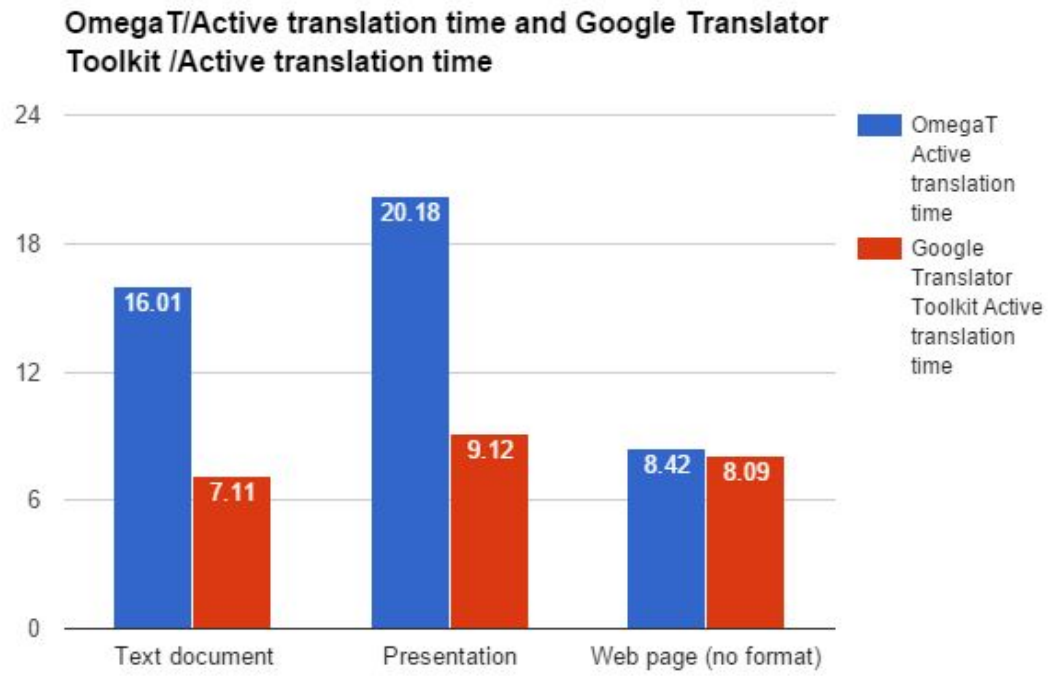

Fig. 1, Active Translation Time difference between OmegaT and Google Translator Toolkit

As stated before, Active Translation Time is referred to the time in which translators are actively engaged in several activities not entirely related to the final stages of the translation process such as revising and proofreading. These moments may include: feeding or uploading the original files to be translated, preparing or selecting the Termbase and Translation Memory to be used in the process, as well as revising the segmentation rules and either join or separate (merge and dissociate) specific segments. In this case we can observe a drastic difference in the times allotted to these activities between two CAT tools. In the the instances where students were requested to translate a text document and a slideshow, both under 200 word in length, Active Translation Time using OmegaT was significantly longer than when using Google Translator Toolkit. This is due to the fact that at least 4 additional steps are required during the feeding process and the verification of the segments involves up to 2 additional steps in the case of OmegaT whereas these processes are automatically

rendered in Google Translation Toolkit. Also, unlike Google Translator Toolkit, OmegaT lacks a default Automatic Translation module which causes students to switch back and forth from typing to verifying tags and other minutiae. There is very little difference between the times taken to translate a simply formatted web page but this may be due not only to the simplicity of the document but also to the rather short length of the document and the ease of the tags that html elements often have. We can also add that by the time the web page has been translated the translation memory in both tools is sufficiently fed with segments in order to render context matches or total matches easily.

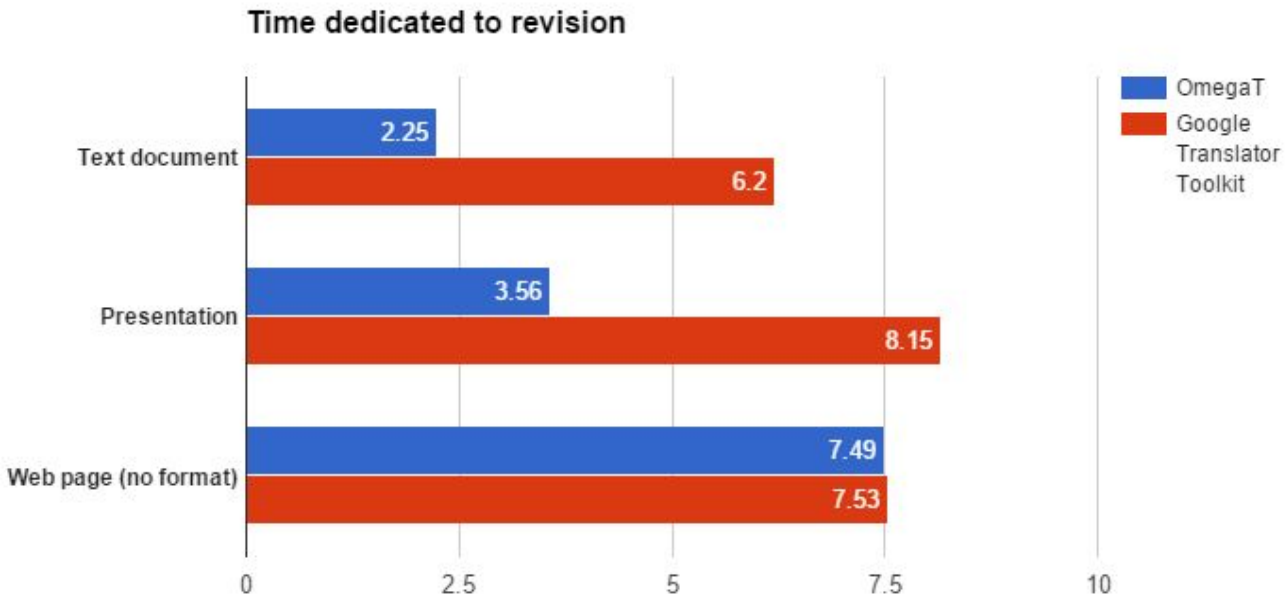

Fig. 2, Difference in Time dedicated to revision between OmegaT and Google Translator Toolkit

The time allocated to revision is differentiated between the CAT tools compared by means of the overall time that is dedicated to each individual stage in the revision process. For example in the case of OmegaT the start of the revision process is clearly drawn due to the fact that the entire translation must be finished first prior starting to revise or locate mistakes in syntax and mistranslations. In the case of Google Translator Toolkit the application automatically renders a machine translated version of the document which reduces considerably the time taken to apply the final revision stages. In the case of translating the web page both Google Translator Toolkit took less time than in the previous processes due to the short length of the document and how identifiable the format tags were. Student Translators seemed to take significantly longer in Google Translator Toolkit due to the fact that most of the actual time is dedicated to revision more than active

translation of the segments. Is very important to notice that the slideshow in this case was heavily formatted and tags were predominant therefore it was observed that a bigger effort was put into retaining the format more than revising for mistakes in translation.

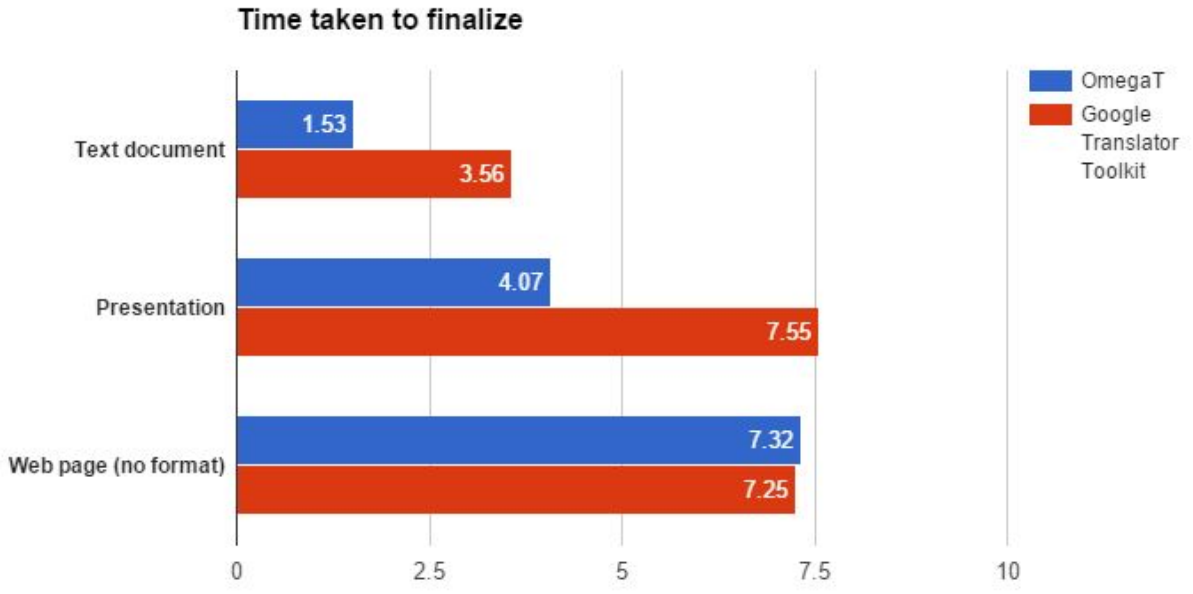

Fig. 3, Difference in Time to finalize the translated documents between OmegaT and Google Translator Toolkit

When we refer to the finalization of a translation project we include the processes related to generation of the final documents and may include: tag verification, Quality assurance and testing of the final file, save to file, print, publish, and comparison of the product to the original or source file. In the graph above we can observe of the times allocated to finalize each document differ greatly depending on the tool that was used and the type of text. The two major differences in time are due to the type of document or source format included at the start of a project and we can follow such differences with how those types of documents behave in each software. Student Translators encountered some visible difficulties upon generating the final documents in saving them in Google Translator Toolkit mostly due to it depending completely on Internet access. At the time of conducting this task the connection presented some issues and Student Translators couldn't effectively download the first two texts. In the case of the third document (the web page), as with the slideshow, tag verification was necessary and therefore Student Translators were required to conduct it which allotted for a longer period of finalization. Upon rendering the translation of the page complete, Student Translators were required to test it making this activity the longest of the

three since no associated web browser were found at the moment of testing the page and they had to associate a browser by means of selecting it as a default application for this type of files.

B. Analysis of data related to mistakes found in the final product or translated texts

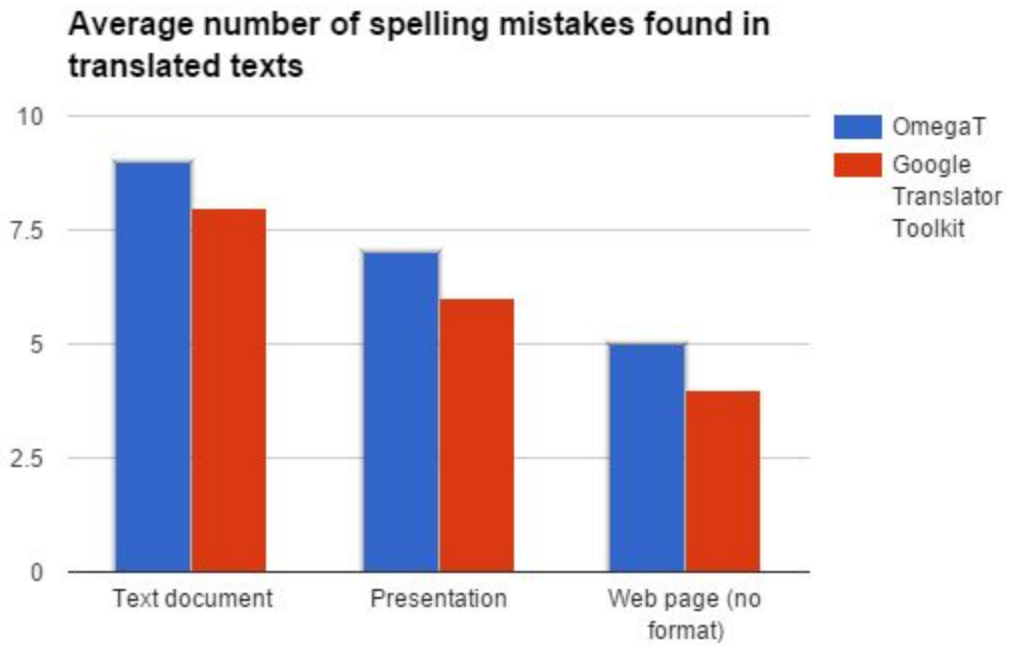

Fig. 4-a, Difference in the number of spelling mistakes in the translated texts between OmegaT and Google Translator Toolkit

	Omega T	Google Translator Toolkit
Text document	9	8
Presentation	7	6
Web page (no format)	5	4

Fig. 4-b, Detail of the number of spelling mistakes in the translated texts in OmegaT and Google Translator Toolkit

The translated documents in OmegaT presented a higher incidence of spelling mistakes with an index that approached the 0.50 incidence level or an average of half a mistake higher than its counterpart, Google Translator Toolkit, per translated document across all categories. This may be due to the fact that OmegaT displays tags in line with the same text which some Students Translators may find somewhat confusing. Another possible reason is due to Google Translator Toolkit being able to render automatically the translated text prior the Student Translator actually delving into the task it actually helped students avoid spelling mistakes since most segments where this sort of mistakes are present where avoided entirely. The low index of spelling mistakes in both Computer Assisted Translation Tools was also lowered due to the fact of adding a revision and proofreading stage.

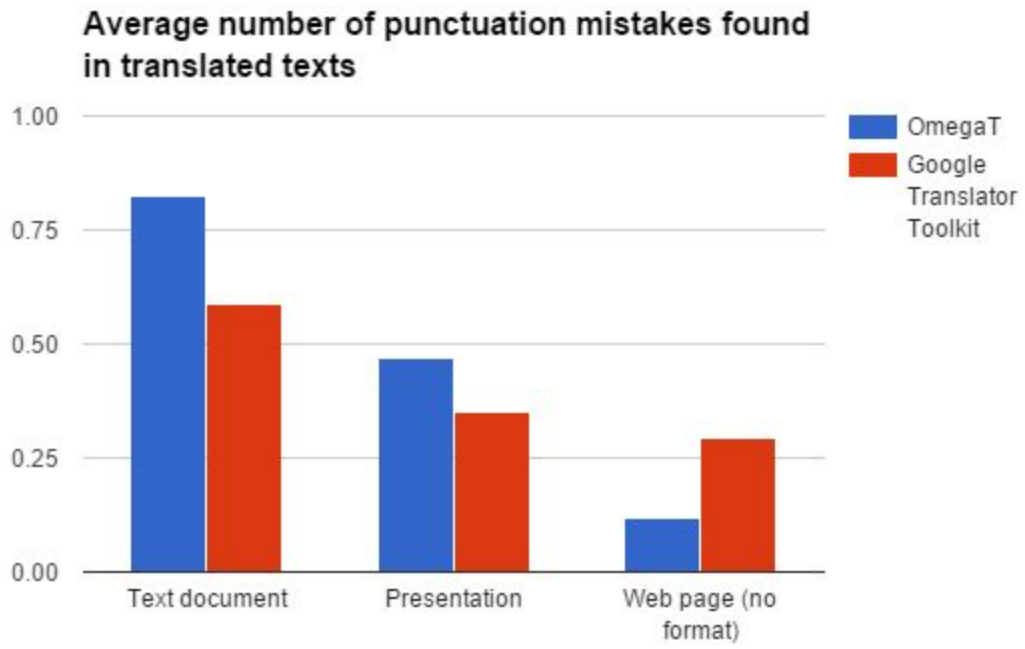

Fig. 5-a, Difference in mistakes in punctuation in the translated texts between OmegaT and Google Translator Toolkit

	OmegaT	Google Translator Toolkit
Text document	14	10
Presentation	8	6
Web page (no format)	2	5

Fig. 5-b, Detail of number of mistakes in punctuation in the translated texts between OmegaT and Google Translator Toolkit

Mistakes in punctuation were distinctively higher in OmegaT when it came to analyzing the text document and presentation. However, There is a drastic shift upon reviewing the results of the translation of the web page setting the margin of difference to more than half of the ones encountered in OmegaT. As we have read before OmegaT relies more on the manual input from the translator and therefore the margin for error is larger when it comes to large sections of text such as the segments encountered in the text document and slideshow but in the case of the web page OmegaT is capable of rendering the tags inline and therefor Student Translators are capable of delimit where the tags start and and in each instance of the segments where they are present more efficiently than Google Translator Toolkit. This difference allowed Student Translators to identify instances of capitalization and periodicity with more ease in OmegaT while translating a web page.

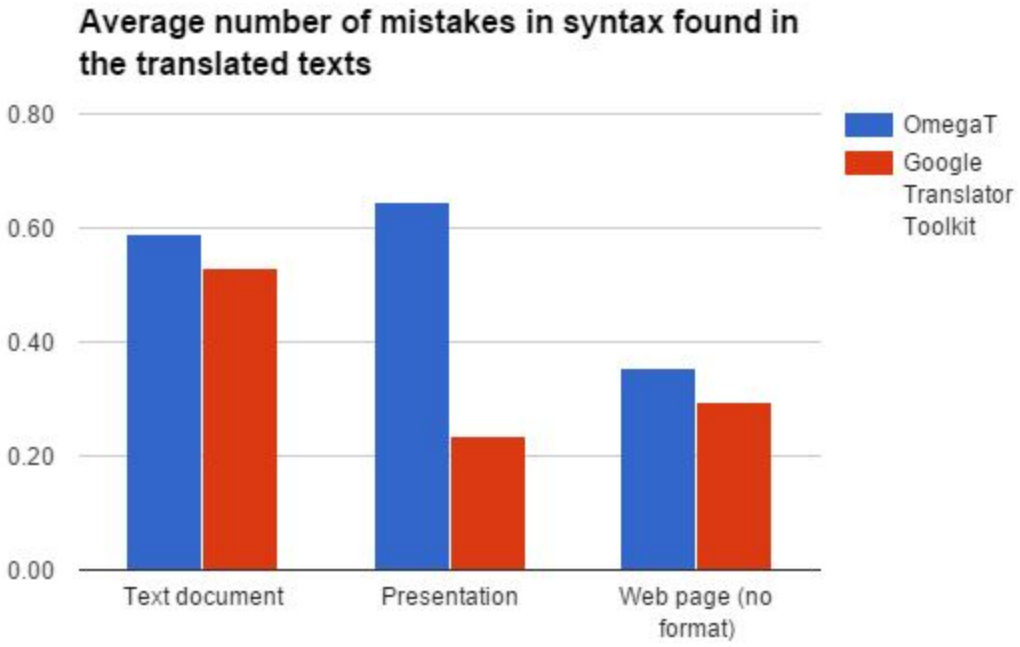

Fig. 5-b, Difference in mistakes in syntax in the translated texts between OmegaT and Google Translator Toolkit

	OmegaT	Google Translator Toolkit
Text document	10	9
Presentation	11	4
Web page (no format)	6	5

Fig. 5-a, Detail of numbers of mistakes in syntax in the translated texts between OmegaT and Google Translator Toolkit

In the case of mistakes in syntax there is very little difference between the index observed in most of the translated texts of Student Translators while using the two proposed CAT tools. The only noticeably large gap between margins of syntax errors occurred in the translation of the slideshow where the margin while using OmegaT reaches almost a 0.7 average number of errors while the translated slideshows rendered in Google Translator Toolkit presented fewer than 0.3 mistakes in syntax. Some conclusions may be drawn from this fact. Errors in syntax were analyzed in terms of Student Translators rendered the translation of simple structures while more complex structures involving multiple nested clauses were omitted for the sake of simplicity. The mistakes in syntax found in the translated documents spanned from mere misplacements of a preposition or verb to incongruities product of lack of knowledge of such structures. The importance of quantifying these sort of mistakes lies on the capabilities that the software has in order to point out such incongruities. Google Translator Toolkit counts with a built in automatic translation module which helps into retaining structural congruity in the target language whereas OmegaT relies heavily on the previous knowledge of the translator and also multiple revisions if possible.

C. Analysis of data related to the levels of difficulty perceived by Student Translators related to the operation of both Computer Assisted Translation Tools

a. Finding the options and controls to create a new project

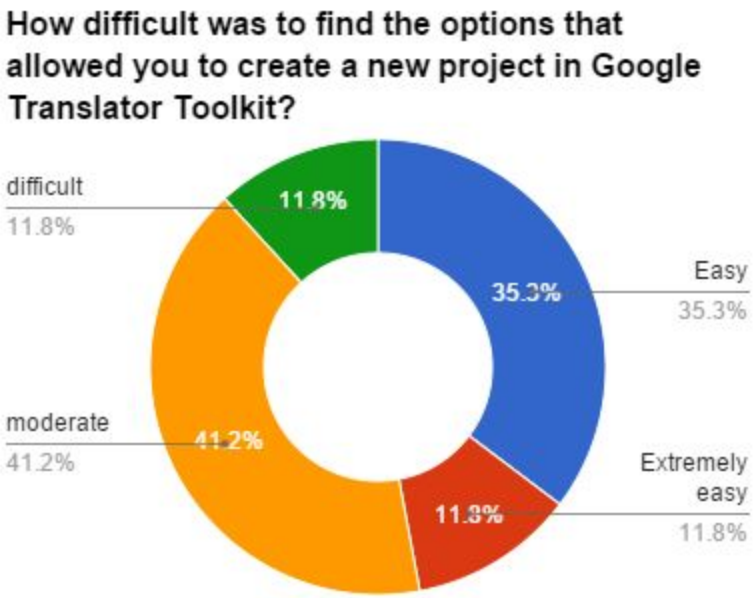

Fig. 6-a, Perceived difficulty of locating the controls that allowed Student Translators to create a new project in Google Translator Toolkit.

Most Student Translators who were surveyed ranked the difficulty level of the task of finding the options that allowed them to create or start a new project in Google Translator Toolkit as moderate and another large percentage of them also ranked it as being easy. As we can see the remaining percentages deemed the distribution and position of elements such as file placement and saving options as difficult or extremely easy. One of the advantages that Google Translator Toolkit offers is that the layout of the interface is very similar to other web tools that Student Translators may already use.

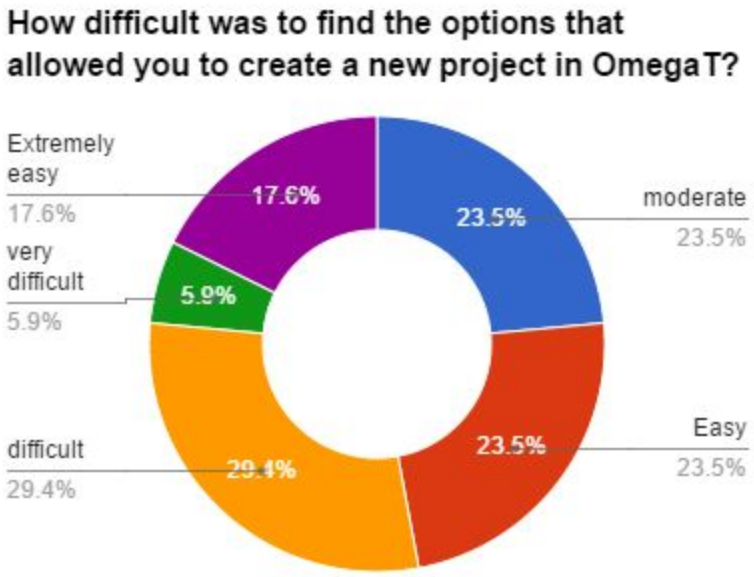

Fig. 6-b, Perceived difficulty of locating the controls that allowed Student Translators to create a new project in OmegaT.

In contrast, a rather large percentage of Student Translators surveyed ranked the level of difficulty in locating the options and controls to start a new translation project in OmegaT in the difficult and moderate range. This may be due to the fact that OmegaT has a particular unique way to create projects which allows translators to locate the source and translated files in different folders within a single folder structure where the rest of the assistive components such as the Translation Memory or TM, and glossary can be located as well. The degree of precision that this task involves makes the overall process more complex if previous knowledge of folder structures and creation of new folders is lacking. To this we may add a rather unique menu and button layout as well as its own distinctive nomenclature of options and user interface labels.

b. Creating a new translation project

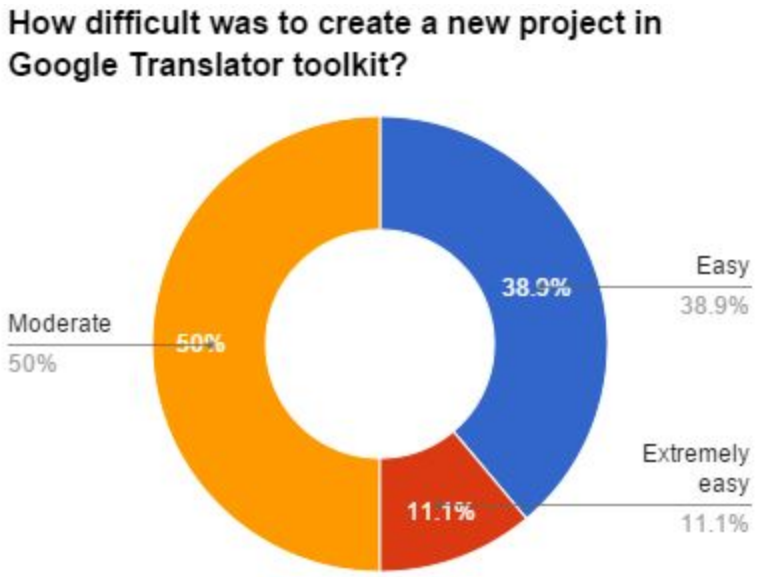

Fig. 7-a, Perceived difficulty upon creating a new project in Google Translator Toolkit

Even though most of Student Translators who were surveyed ranked the process of creating an entire new project from within Google Translator Toolkit across the "easy spectrum" what is truly remarkable is that these numbers differ from the previous results destined to show how the degree of difficulty was perceived in relation to the task of merely locating the controls that would enable Student Translators to effect this task. In this case Student Translators perceived a low degree of difficulty upon actually creating a new project a task that would involve in the case of Google Translator Toolkit to log into their private accounts, name the project, upload the files and choosing whether to use a third party machine translation server or their own. The biggest difference that we encountered was the fact that Google Translator Toolkit automatically chooses the default TM or Translation Memory thus preventing Student Translators of breaking the continuity of the process or backtrack in order to confirm that the adequate choices were made from the start.

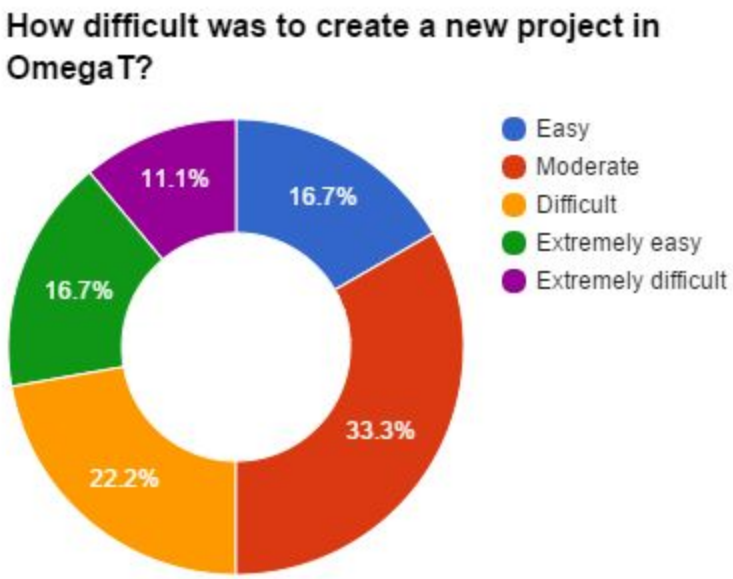

Fig. 7-b, Perceived difficulty upon creating a new project in OmegaT

In the case of OmegaT, the Student Translators Surveyed ranked the degree of difficulty of creating a new project as either difficult or moderate. This may be due to OmegaT requiring the user to specify destination folders for the project, which must be empty prior starting, and a manual configuration of the Translation Memory module, which although rather simple and straightforward it may still be a little daunting to configure for the inexperienced user. The layout of the main options can also lead to some confusion if not followed properly.

c. Loading files to translate

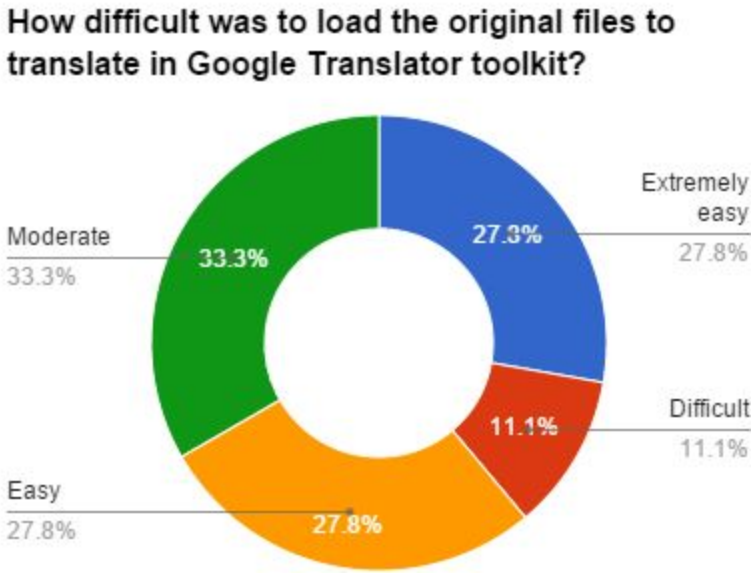

Fig. 8-a, Perceived difficulty upon loading raw or untranslated documents into a new project in Google Translator Toolkit

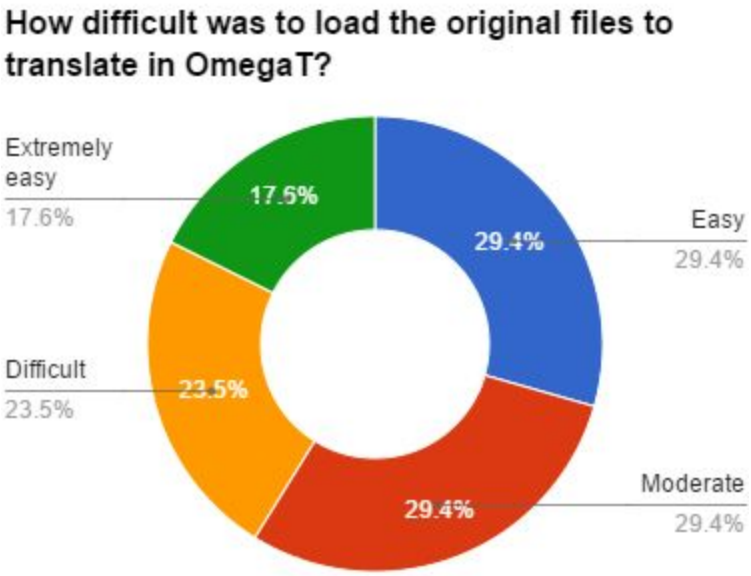

Fig. 8-b, Perceived difficulty upon loading raw or untranslated documents into a new project in OmegaT

The difference between the perceived difficulty upon loading raw or source files into both CAT tools is noticeable. While most Student Translators ranked Google Translator Toolkit as easy or moderate OmegaT received a rank of higher difficulty which may be due to the fact that additional steps are required such as conversion of the source files into Open Source formats in order to be accepted by the file loader whereas Google Translator Toolkit allows this process to be more streamlined and straightforward .

d. Locating the controls to start translating the document after loading it

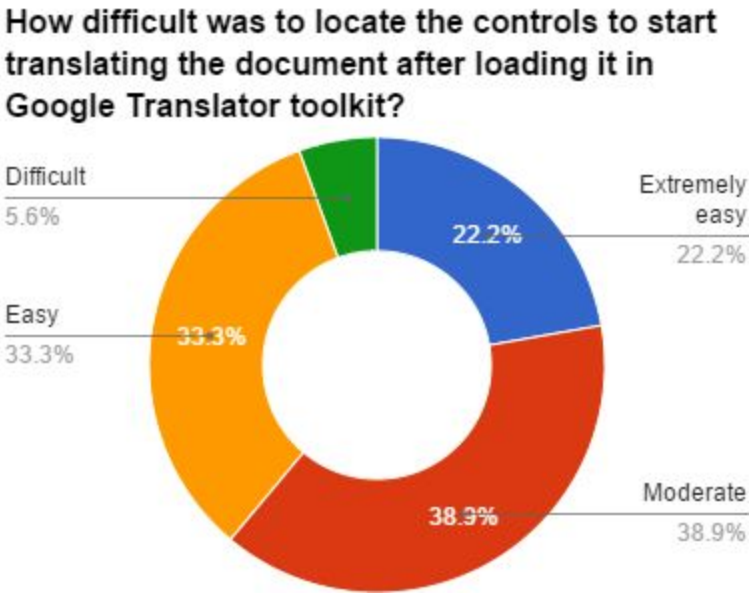

Fig. 9-a, Perceived difficulty when locating the controls to start translating a new document after loading it in Google Translator Toolkit

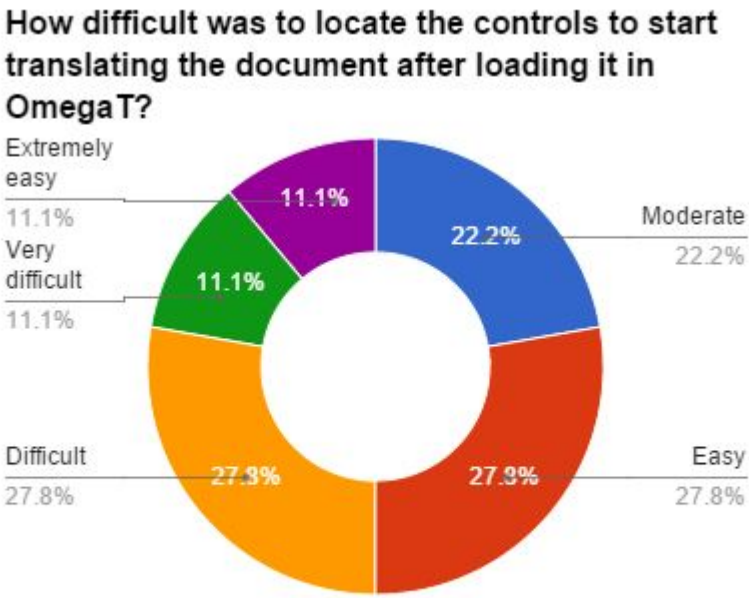

Fig. 9-a, Perceived difficulty when locating the controls to start translating a new document after loading it in Google Translator Toolkit

The difference between the perceived difficulty upon locating the controls that allowed Student Translators to start translating the texts is again very noticeable placing OmegaT in a higher level. The layout of the tool may be a significant and recurrent component of the layout.

e. Saving the progress of the translation

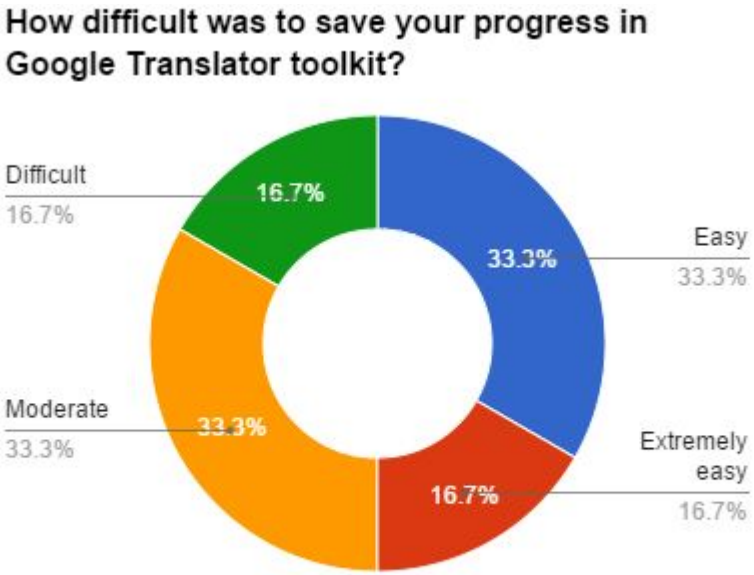

Fig. 10-a, Perceived difficulty upon saving a project in Google Translator Toolkit

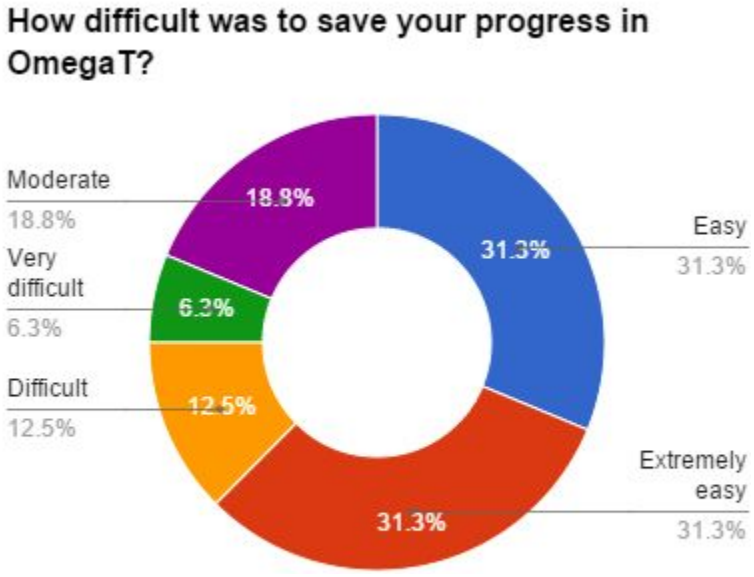

Fig. 10-b, Perceived difficulty upon saving a project in OmegaT

Upon saving the changes in the target texts as they are translating, Student Translators conducting their practice in Google Translator Toolkit perceived a higher level of difficulty since the layout of the software despite being fairly simple has a disabled automatic save option which may confuse Student Translators as they progress whereas in OmegaT Student Translators successfully committed changes to the page as they progressed without much difficulty.

f. Identifying new matches

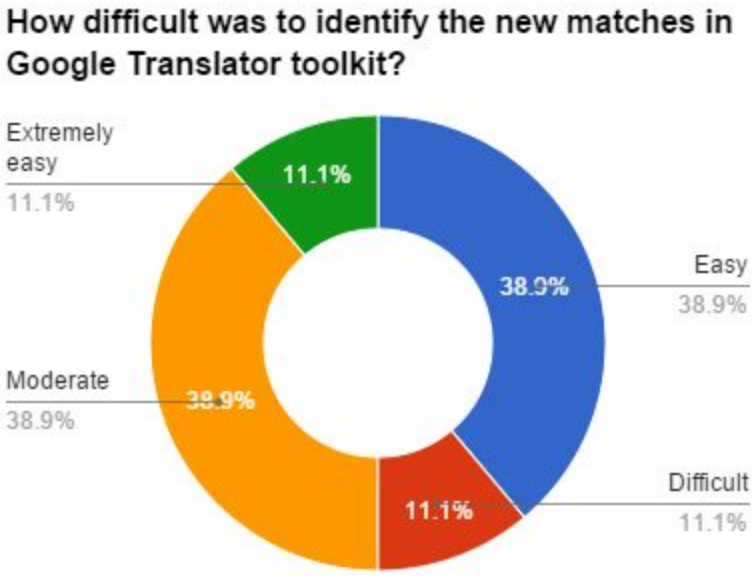

Fig. 11-a, Perceived difficulty upon identifying new matches in Google Translator Toolkit

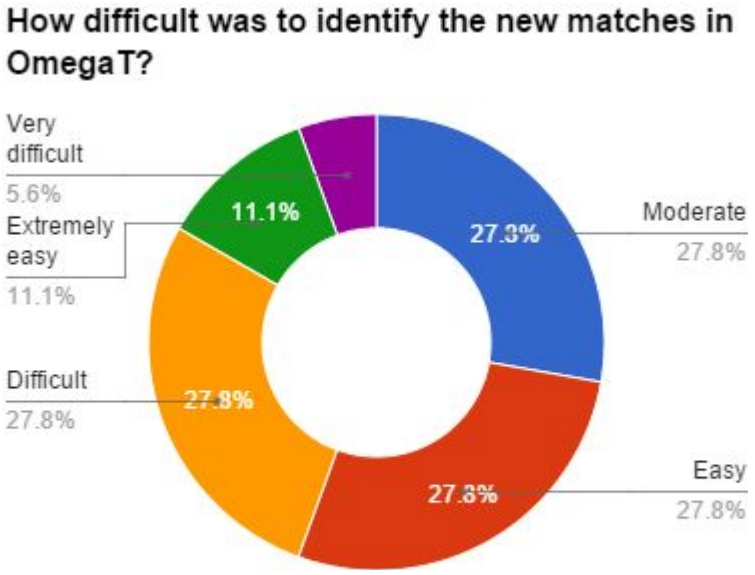

Fig. 11-a, Perceived difficulty upon identifying new matches in Google Translator Toolkit

The degree of perceived difficulty for this instance describes the availability of the upcoming matches as Student Translators advance in the translation of each document. This indicator in specific is related to how much of the Translation Memory or TM the Student Translators can use at a specific moment. In the case of OmegaT the perceived difficulty is higher may be due to the fact that matches appear in a different module and the percentage of concordance is shown as an estimate rather than the actual concordance index whereas in Google Translator Toolkit matches are immediately replaced upon being identified.

g. Identifying fuzzy matches and total matches

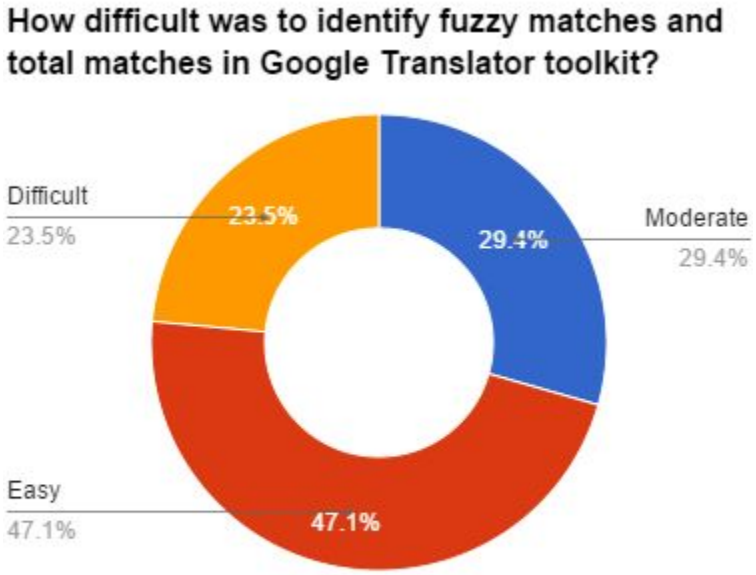

Fig. 12-a, Perceived difficulty upon identifying fuzzy matches and total matches in Google Translator Toolkit

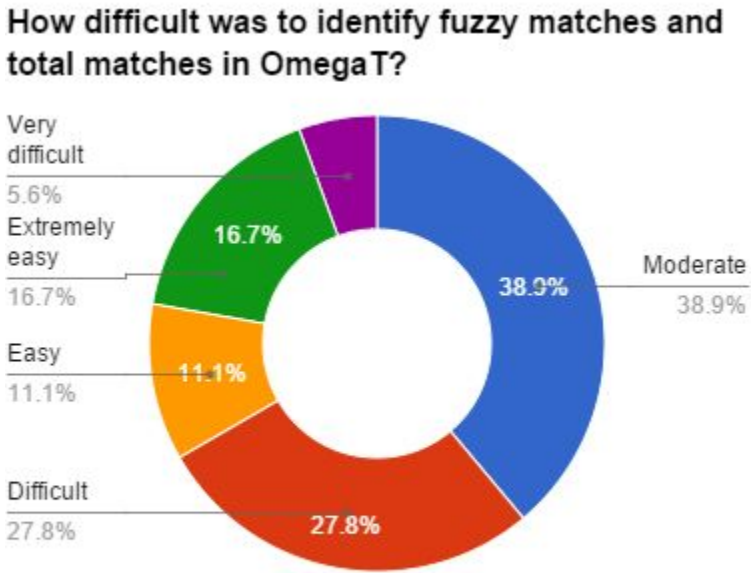

Fig. 12-b, Perceived difficulty upon identifying fuzzy matches and total matches in OmegaT

In the case of identifying fuzzy or partial matches and complete or total matches the difficulty level perceived by most of Student Translators surveyed was similar across both CAT tools varying relatively little. This may be due to both CAT tools having similar features and options when it comes to identifying new fuzzy matches such as a selector for the options turned by the Translation Memory.

h. Use of terminology support (if any) while translating

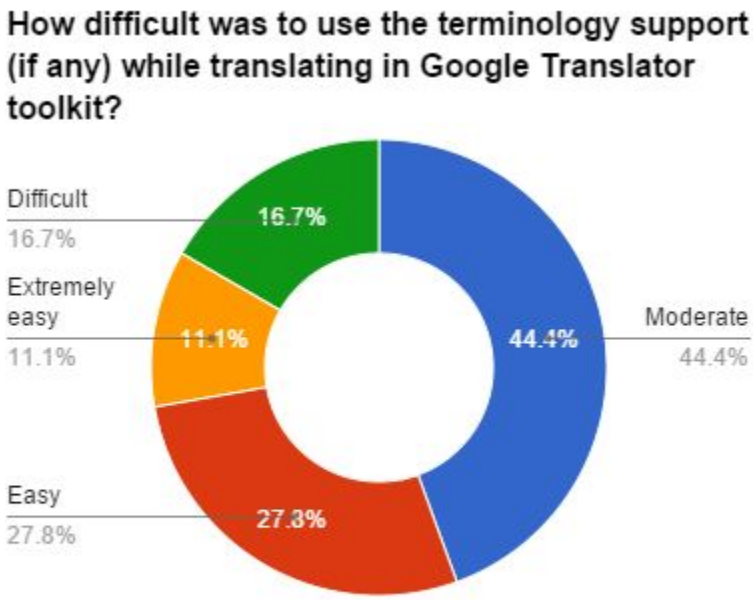

Fig. 13-a, Perceived difficulty upon using terminology support in Google Translator Toolkit

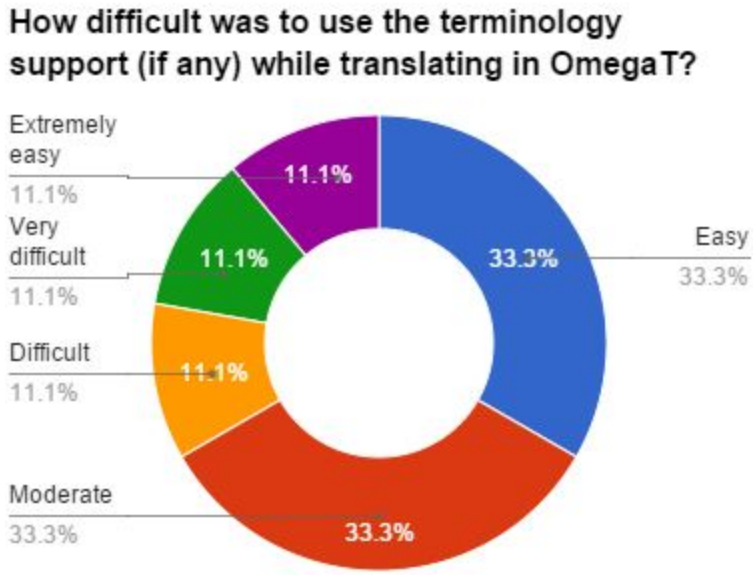

Fig. 13-b, Perceived difficulty upon upon using terminology support in OmegaT

Both CAT tools were similarly placed according to their degree of difficulty when it comes to use of terminology support since both allow the use of user generated glossaries or Terminology Databases at a very simple and basic level. Student Translators were able to grasp the initial concepts of suggestion dictionaries and user defined glossaries without much effort.

i. Revising and proofreading the translated texts

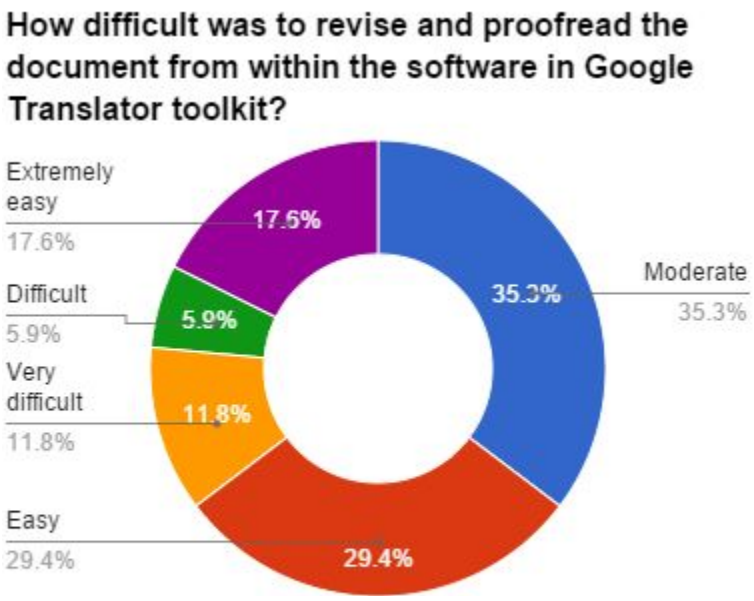

Fig. 14-a, Perceived difficulty upon revising and proofreading the documents from within the software interface in Google Translator Toolkit

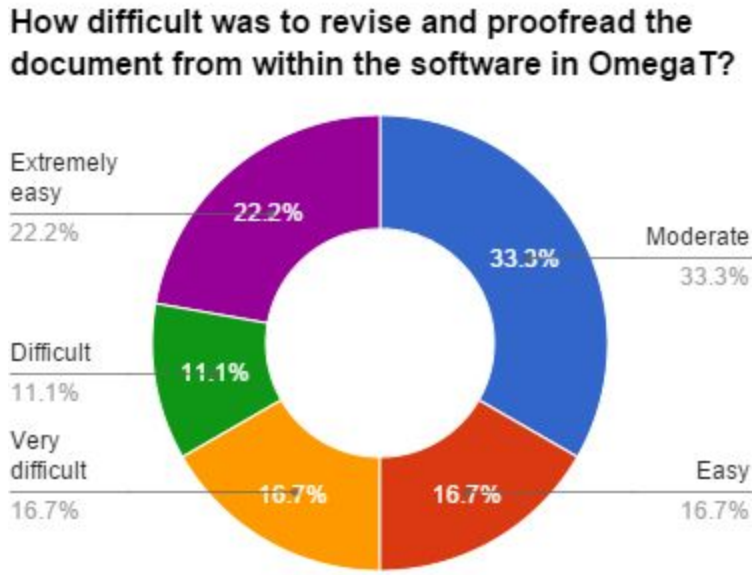

Fig. 14-b, Perceived difficulty upon revising and proofreading the documents from within the software interface in Google Translator Toolkit

There is very little difference between the perceived difficulty between both CAT tools when it comes to revising and proofreading from within their interfaces. Most of Student Translators surveyed ranked the difficulty level of the proofreading and revising options within the CAT tools as low or moderate and allowed them to easily find and correct mistakes.

j. Navigating through the controls to add, remove or modify the format of the document

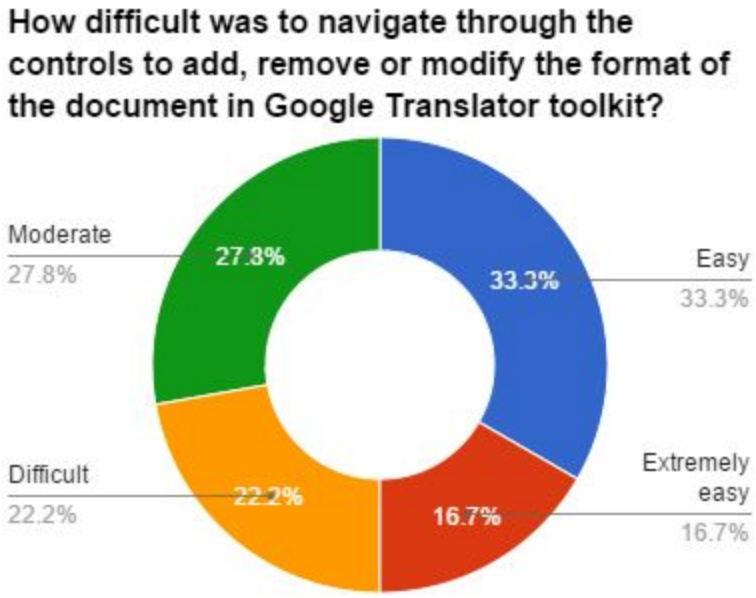

Fig. 15-a, Perceived difficulty upon navigating through the main controls to carry out editing operations in Google Translator Toolkit

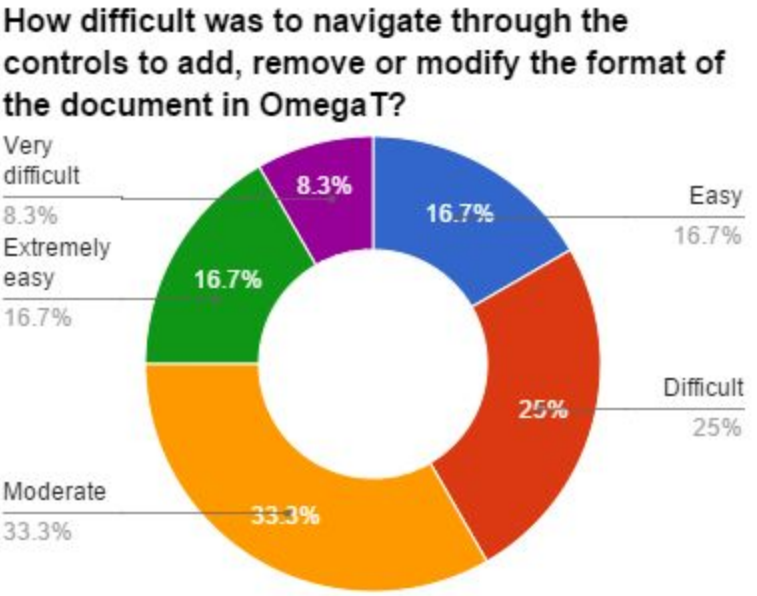

Fig. 15-b, Perceived difficulty upon navigating through the main controls to carry out editing operations in Google Translator Toolkit

The perceived difficulty in moving through the controls that allowed to format the text and final leans towards OmegaT since the layout is not s intuitive as possible.

k. Generating the final documents

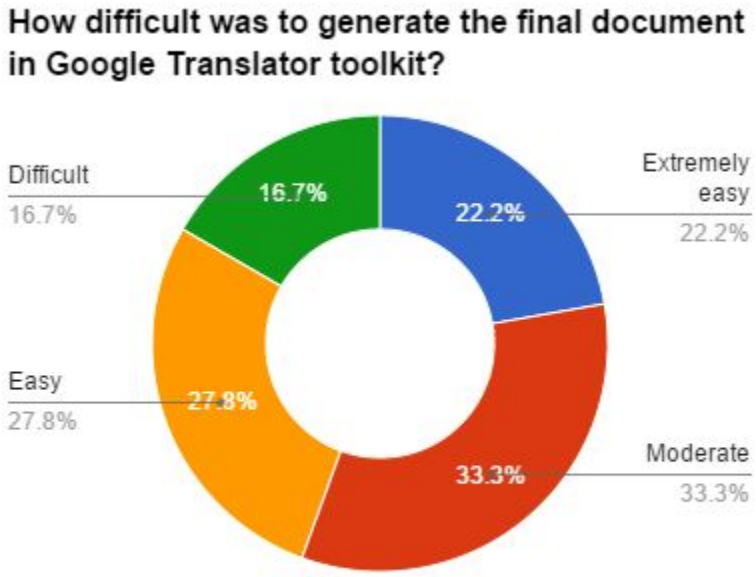

Fig. 16-a, Perceived difficulty generating the final documents or translated texts in Google Translator Toolkit

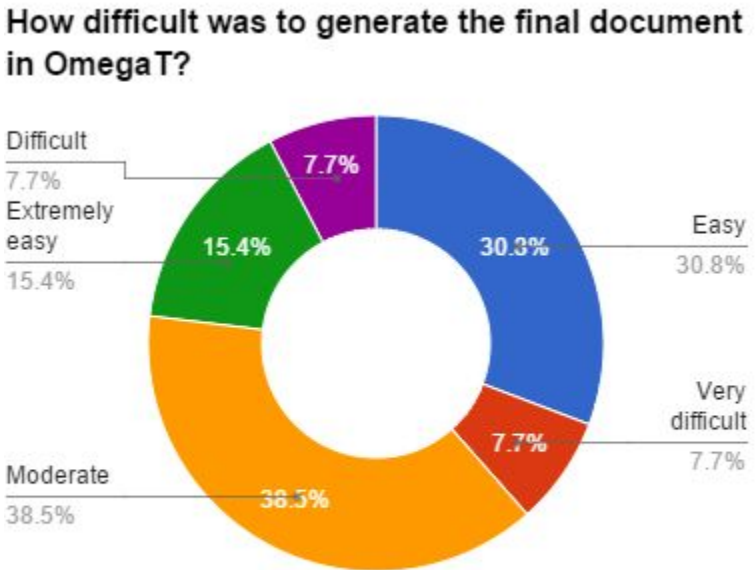

Fig. 16-b, Perceived difficulty generating the final documents or translated texts in OmegaT

In both scenarios Student Translators ranked these CAT tools within the range from easy to moderate, most Student Translators did not have as many issues with the generation, format, and publication (either print or in digital).

l. Working with text documents

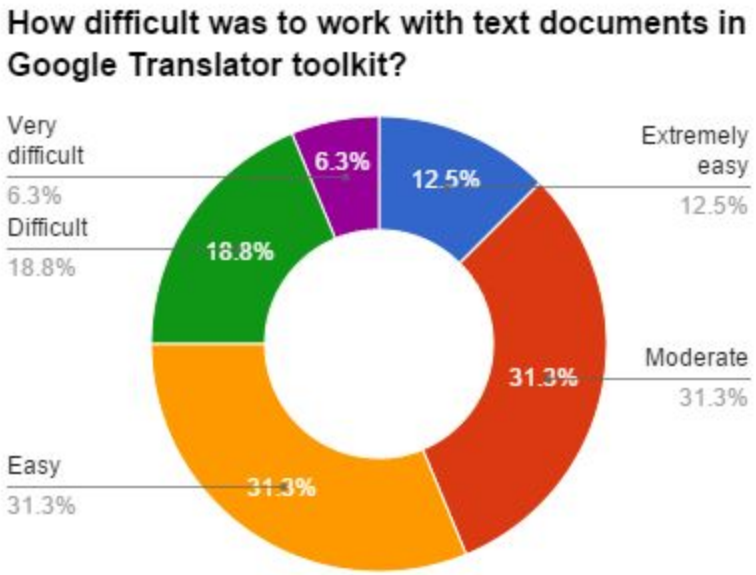

Fig. 17-a, Perceived difficulty upon translating text documents in Google Translator Toolkit

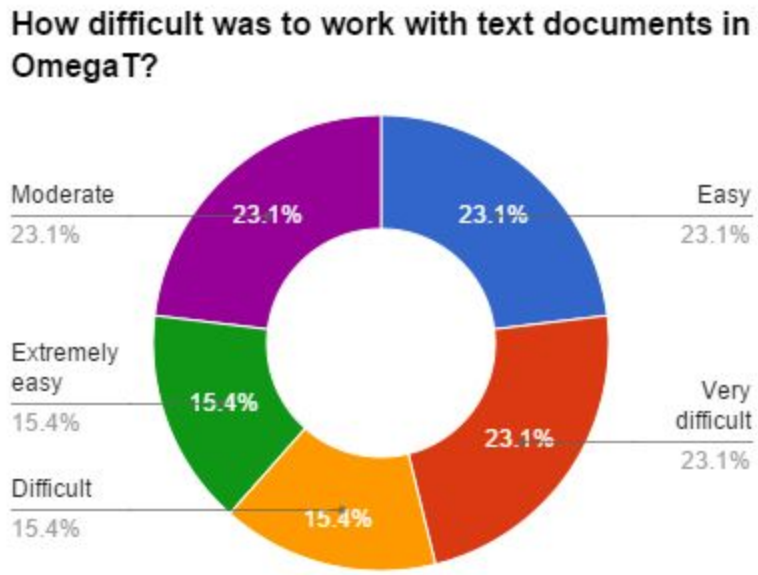

Fig. 17-b, Perceived difficulty upon translating text documents in OmegaT

In this case there is an obvious difference in the perceived difficulty when working or translating text documents which may include a wide range of extension formats such as .doc or .docx or even .odt which could give us an inlink into why the perceived difficulty is higher when it comes to text documents since in order to work in OmegaT documents have to be saved to a compatible file extension format such as .odt.

m. Work with with slideshows

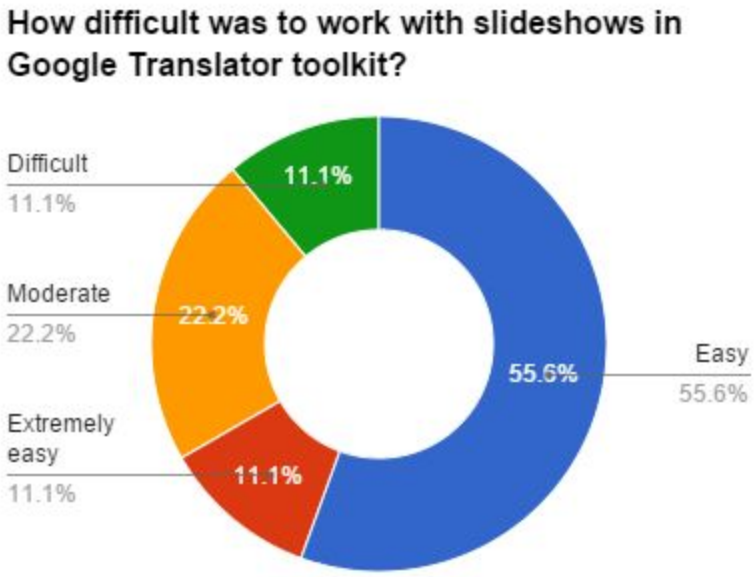

Fig. 17-a, Perceived difficulty upon working with slideshows in Google Translator Toolkit

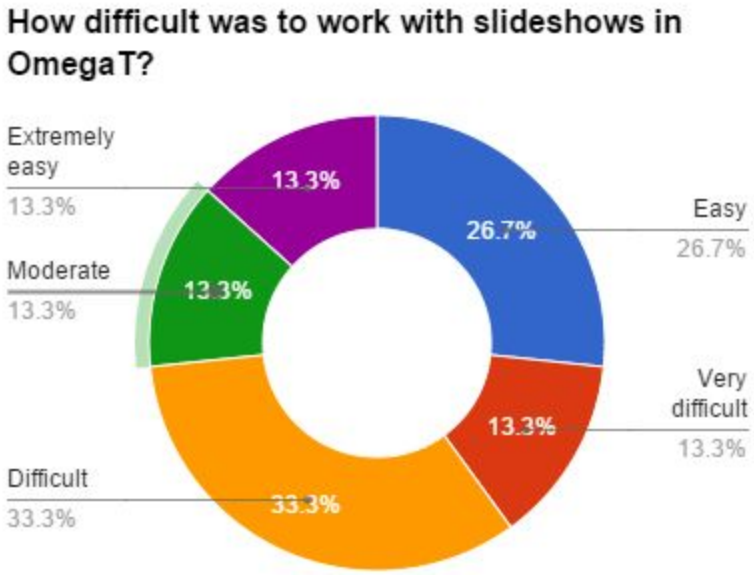

Fig. 17-b, Fig. 17-a, Perceived difficulty upon working with slideshows in OmegaT

The perceived difficulty in terms of working or translating slideshows suffers from the same hinderance described above, the lack of compatibility between proprietary extension formats and the interface of OmegaT. This lack of compatibility is a recurrent descriptor among CAT tools.

n. Work with web pages

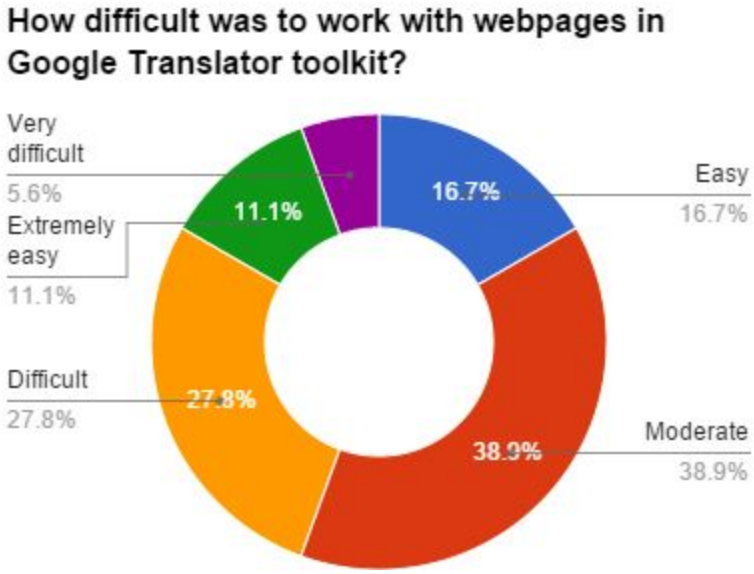

Fig. 18-a, Perceived difficulty upon working with Web Pages in Google Translator Toolkit

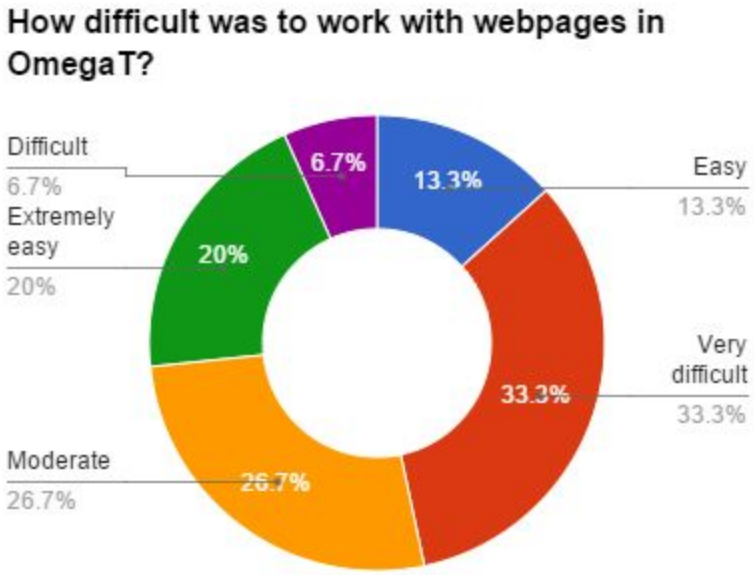

Fig. 18-b, Perceived difficulty upon working with Web Pages in Omega T

Once again, the perceived difficulty levels while translating a webpage varied greatly leaning towards the very difficult in the case of OmegaT. At a glance we could argue that the same difficulty was found upon loading the file but due to observation we may describe the process in which the file was not converted or saved in a different form at any point. The difficulty may lie in another component such as tag verification or translation memory which needs to be correlated with the previous statistics.

XI. Analysis

A. Analysis of differences between times of activities in the translation process

The main differences between the results obtained from the observed time dedicated by Student Translators to the different aforementioned activities in the individual result analysis section are summarized below.

	OmegaT	Google Translator Toolkit	Conclusion drawn from observed practices correlated to observed results
Active Translation Time	It was considerably higher while translating text documents and slideshow presentations but remained close to its competitions while translating a simple unformatted webpage	It was lower by more than half than its competitor with the exception of translating web pages were its leading position remained by a very small margin (33 second to be exact)	Student Translators work time favored Google Translator Toolkit due to its built- in machine translation module which integrates with the official repository of translation matches from Google.
Level of difficulty perceived by Student Translators (based on the following indicators: ● Finding the options and controls to create a new project ● Creating a new translation project ● Loading files to translate ● Locating the controls to start translating the document after loading it ● Saving the progress of the translation ● Identifying new matches ● Identifying fuzzy matches and total matches ● Use of terminology support (if any) while translating)	The level of difficulty perceived by Student Translators while attempting the activities related to the first stages of the translation process in this tool was deemed uniformly as either difficult or moderate. The identification process of potential matches and context matches was considered as being more difficult in this tool.	The level of difficulty perceived by Student Translators while attempting the activities related to the first stages of the translation process in this tool was deemed as either easy or moderate including the processes to identify potential matches and context matches.	In terms of perceived difficulty Google Translator Toolkit remained in the lead between the two compared CAT tools. The surveyed Student Translators deemed the accessibility to controls as well as the connection of processes as being less difficult to both access and operate when compared to OmegaT. we could observe that some of the decisions in the design of the User Interface played an important factor in the choice of most Student Translators surveyed.
Time dedicated to revision	Time for revision was much lower in OmegaT than in Google Translator Toolkit however the revision of the unformatted web	In the case of the time allocated for revision, Students Translators took longer in verifying and correcting mistakes	Upon revision of the observation notes, some of the most relevant indicators to point out are: the capabilities of the software related to automatization of retention of source format and the integration of revision modules

	page remained very close to the results obtained by its competitor	in Google Translator Toolkit than in OmegaT	pertaining terminology management, and concordance of the overall translation.
Level of difficulty perceived by Student Translators (based on the following indicators: ● Revising and proofreading the translated texts ● Navigating through the controls to add, remove or modify the format of the document)	In the case of the levels of difficulty related to related to proofreading and revising OmegaT was on par with Google Translator Toolkit but this was not the case of the perceived difficulty upon navigating through the controls to add, remove or modify the format of the document from within the actual native interface of the application.	In the case of the levels of difficulty related to related to proofreading and revising Google Translator Toolkit received similar appraisals from Students Translators surveyed but received higher index points in terms of accessibility and ease of use upon navigating through the controls to add, remove or modify the format of the document from within the actual native interface of the application.	Upon reviewing observation notes some of the most relevant points that helped identify the possible underlying causes for the perceived difficulty of using the software during the revision, proofreading, and formatting of a document are: location of the controls in terms of providing ease of access to the most common tasks related to format and proofreading of the final document. Automatization of the revision and proofreading processes by means of control allocation.
Time taken to finalize the translated texts	The time taken to finalize documents in OmegaT was significantly shorter during the final steps observed during the tests with the exception of the translation project of the web page. Student Translators had to complete less steps in order to process the final document in contrast to the processes found in Google Translator Toolkit.	The time taken by Student Translators in order to finalize the files to translate in Google Translator Toolkit was longer during the first two tasks while remaining longer by 13 seconds in the final stages of the translation process of the web page.	The indicators to follow in order to identify the possible causes related to lengthening the time taken to finalize the documents are: the speed of the automatized process of generating the final documents, the ease of access to the controls that allow generating the final documents, and the perhaps the most important aspect is the integration of an automated machine translation module by default in the tool.
Level of difficulty perceived by Student Translators (based on the following indicators: ● Generating the final documents)	Student Translators perceived both tools to be, in terms of ease of use, accessibility, and how intuitive they are, between the easy and moderate spectrum with very few Student Translators considering it difficult.		the most relevant items to discern in relation to the difficulty level perceived by student translators upon generating the final documents are: the location and distribution of the controls that allow users to effectively generate the final version of the documents prior printing, the overall speed of the process to generate the final documents, and the level of complexity of the formatting options in generating the final documents.

2. Analysis of differences between number of errors in the final version of the translated documents

	OmegaT	Google Translator Toolkit	Conclusion drawn from observed practices correlated to observed results
Average number of spelling mistakes	There is a higher incidence of mistakes in spelling across all the documents that were rendered in OmegaT.	A distinctively lower rate of spelling mistakes were found across all the translated texts.	The indicators that may be related to the incidence of mistakes in spelling in both tools are: the automatization of a machine translated version of the file, the identification of tags in the editor window or editing area of the software may it be supported natively or not.
Average number of mistakes in punctuation	Distinctively higher number of average mistakes in punctuation.	There is a drastic shift in the translation of the web page setting the margin of difference to more than half of the ones encountered in OmegaT.	Some of the observed indicators that may allow for the difference in number of mistakes in punctuation are: automatization of machine translation, retention of format tags in html documents, and segmentation rules.
Average number of mistakes in syntax	The only noticeably large gap between margins of syntax errors occurred in the translation of the slideshow where the margin while using OmegaT reaches almost a 0.7 average number of errors while the translated slideshows rendered in Google Translator Toolkit presented fewer than 0.3 mistakes in syntax.		Upon reviewing the observation notes the problems encountered in terms of misplacement of more complex grammatical structures may have been due to the some of the following indicators: retention of tags in format, lack of concordance module, and distribution of controls used to proofread the overall translation of the document, as well as the lack of knowledge of more advanced strategies in Student Translators.

XII. Conclusions

After analyzing the data obtained from the observed translation sessions and the data gathered by the survey related to the perceived difficulty on very specific elements of the complexity of the user interface which arises from the tasks to complete at various stages as well as the number of mistakes found in the translated texts we conclude the following:

- The user interface of OmegaT although familiar to most experienced users also rendered some tasks more difficult due to the complexity of navigating through the options such as the ones used to create projects and upload documents to be translated.

- The lack of an automated machine translation module by default became a critical component for Student Translators upon qualifying the degree of difficulty related to very specific tasks such as pretranslatting and recognizing large segments in OmegaT

- After reviewing and analyzing the indicators encompassing the opinion of Students Translators who participated in this study we conclude that most Student Translators favored Google Translator Toolkit in terms of usability, adaptability, and how intuitive they perceived the interface and other options available in the tools to be in contrast with OmegaT.

- Although most Student Translators Favored Google Translator Toolkit in terms of perceived difficulty, some of the indicators related to time and number of mistakes also showed a perceivable advantage of using OmegaT. This is due to its reliance on human input, making it a valuable alternative when it comes to the field of Translation Studies. Most Student Translators were able to identify their mistakes prior continuing to further stages in the translation process reducing the time spent

in revision without realizing that the revision stage had been completed before. OmegaT is very valuable since it allows participants to become fully immersed in the translation process which encompasses the earliest stages as well as revision and proofreading.

- Several characteristics of Google Translator Toolkit were further explored by Student Translators due to their ease to access and how intuitive they proved to be.

XIII. Recommendations

After analyzing the results and reaching concluding thoughts on the possible causes to the observed phenomena we recommend the following:

- In terms of User Interface both OmegaT and Google Translator Toolkit need to improve in very specific areas in order to simplify the process of translation for Students Translators as well as to streamline some of the features involved in initiating the process. Some of the shared recommended areas to improve are:
  - The location of controls to manipulate the results displayed by the Translation Memory by placing them under the window presenting the percentage of matches between identified segments and stored results.
  - The distribution of some options within the menus of both tools caused Student Translators to become confused and sometimes to forget the location of important controls. The interface will benefit from a band in which the most commonly used controls can be found and clicked.
  - Both interfaces could benefit greatly from adding easily identifiable icons in a band below the menus containing the most used options to start a project, add files, and maybe easier access to options that help identify matches within a selected Translation Memory.

- In regards to the User Interface of OmegaT and after analysing the data related to the average number of mistakes, time taken to complete the three mains stages of the translation process, and the difficulty levels perceived by Student Translators upon using the software as well as taking into consideration the conclusions drawn above we can recommend:
  - The interface of the menu could be changed to contain a single bar with icons that allow users to access individual commands more easily without having to open drop down menus.

- To include an additional pane that allow users to identify matches more easily, displaying information such as optional matches or fuzzy matches.
- To highlight the segment to which the term displayed in the glossary pane corresponds to.
- To add an automatic machine translation module.

Bibliography

Alcina, A. (2001). *Translation Technologies : Scope, Tools and resources*. International Journal of Translation Studies, (68), 1–33.

Anderman, G., & Rogers, M. (2003). *Translation Today: Trends and Perspectives* (p. 243). The British Library, Multilingual Matters.

Bowker, L. & Fisher, D. (2010). *Computer-Aided Translation, Handbook of translation studies Vol. 1*. Amsterdam / Philadelphia: John Benjamins Publishing Company. 2006, págs. 60-64

Cronin, M. (2013). *Translation in the Digital Age*. Routledge.

EAGLES, Experts, Advisory Group on Language Engineering Standards (1995) *Evaluation of Natural Language Processing Systems.* Final Report, University of Geneva.

Esselink,B. (2000). *A practical guide to localization* Amsterdam / Philadelphia: John Benjamins Publishing Company.

Feder, M. (2003). *Machine - assisted human translation : Its position,* Perspectives : Studies in Translatology, Volume 11, issue 2.

García, I., & Stevenson, V. (2009). Reviews-google translator toolkit.*Multilingual computing & technology*, *20*

Gambier, Y., & Dorrslaer, L. (2010). *Handbook of translation studies Vol. 1*. Amsterdam / Philadelphia: John Benjamins Publishing Company.

Shreve, G. M. (2000). "*Translation at the millennium: Prospects for the evolution of a Profession"*. P.A. Schmidt.

Hayward, M.(2003). B*ecoming a Translator, Second Edition*. Routledge

Mckay, C. (2004). *FREE AND OPEN SOURCE SOFTWARE FOR TRANSLATORS* Panace@: Revista de Medicina, Lenguaje y Traducción, Vol. 7, Nº. 23, 2006, págs. 95-98

Palacz, B. (2003). *A Comparative Study of CAT tools (MAHT workbenches) with Translation Memory Components*. Adam Mickiewicz University.
Retrieved from http://www.transsoft.seo.pl/other/TranslatorTools/TMTChpt1.pdf

Robinson, D. (1997.). *Becoming a Translator*: *An introduction to the Theory and Practice of Translation, Second Edition.* Routledge.

Villianatou, F. (2014). *CAT Tools and Productivity : tracking words and hours.* Translation Journal, Volume 9, No 4, pags 1–6.
Retrieved from http://accurapid.com/journal/34CAT.htm

Williams, j. & Chesterman, A. T*he Map: A Beginner's Guide to Doing Research in Translation Studies.* Routledge, 2014.

Annexes

GLOSSARY

https://docs.google.com/document/d/1wZ6_Eh9ZN4apLrrksSVwiV35dHFyCnx4Iou2pWOXojI/edit?usp=sharing

- **CAT (Computer Assisted Translation) tools:**
    - A CAT tool is a computer program that helps to translate text documents more efficiently through four main functions: A CAT tool segments the text to be translated in segments (sentences) and presents the segments in a convenient way, to make translating easier and faster
- **Terminology Management System:**
    - A type of translation software that enables users to efficiently collect, process, and present terminology.
- **Translation memory:**
    - A translation memory, or TM, is a database that stores "segments", which can be sentences, paragraphs or sentence-like units (headings, titles or elements in a list) that have previously been translated, in order to aid human translators.
- **Open source software:**
    - Open source software is software that can be freely used, changed, and shared (in modified or unmodified form) by anyone. Open source software is made by many people, and distributed under licenses that comply with the Open Source Definition.
- **Machine translation:**
    - Translation procedures carried out by a computer
- **Fuzzy matching:**
    - Fuzzy matching is a technique used in computer-assisted translation as a special case of record linkage. It works with matches that may be less than 100% perfect when finding correspondences between segments of a text and entries in a database of previous translations.
- **Front end**
- **back end**
    - Front-end and back-end are terms used to characterize program interfaces and services relative to the initial user of these interfaces and services. (The

"user" may be a human being or a program.) A "front-end" application is one that application users interact with directly. A "back-end" application or program serves indirectly in support of the front-end services, usually by being closer to the required resource or having the capability to communicate with the required resource. The back-end application may interact directly with the front-end or, perhaps more typically, is a program called from an intermediate program that mediates front-end and back-end activities.

- **GNU**
- **GPL**
  - The GNU General Public License (GNU GPL or GPL) is a widely used free software license, which guarantees end users (individuals, organizations, companies) the freedoms to run, study, share (copy), and modify the software. Software that allows these rights is called free software and, if the software is copylefted, requires those rights to be retained. The GPL demands both. The license was originally written by Richard Stallman of the Free Software Foundation (FSF) for the GNU Project.

- **Database**
  - A database is an organized collection of data. It is the collection of schemas, tables, queries, reports, views and other objects. The data are typically organized to model aspects of reality in a way that supports processes requiring information, such as modelling the availability of rooms in hotels in a way that supports finding a hotel with vacancies.

- **Parallel Corpus**
  - A translation Memory

- **Scenario testing**
  - Scenario testing is a software testing activity that uses scenarios: hypothetical stories to help the tester work through a complex problem or test system. The ideal scenario test is a credible, complex, compelling or

motivating story the outcome of which is easy to evaluate. These tests are usually different from test cases in that test cases are single steps whereas scenarios cover a number of steps.

- **Black-box testing**
  - Black-box testing is a method of software testing that examines the functionality of an application without peering into its internal structures or workings. This method of test can be applied to virtually every level of software testing: unit, integration, system and acceptance. It typically comprises most if not all higher level testing, but can also dominate unit testing as well.

Links to live surveys and results:

survey omegaT: http://goo.gl/forms/jAOtVtv3IF

survey on issues found in OmegaT:
http://goo.gl/forms/chw0saee0n

survey on Google Translator Toolkit;
http://goo.gl/forms/w0Onzyxqe5

survey on issues found in Google Translator Toolkit:
http://goo.gl/forms/E71BpN2nAv

Instruments and results for data analysis :
https://docs.google.com/spreadsheets/d/189JWccDdS2N_--6jyvsb953nkDFacJRCPbH1U_IU1-M/edit?usp=sharing

Online stopwatch used by students:

http://www.online-stopwatch.com/large-stopwatch/

Survey on CAT tools - OmegaT

On a scale from 1 - 5 (difficulty) answer the following questions related to the ease of use of OMEGAT translation software

1. **How easy was to find the options that allowed you to create a new project?**
   Mark only one oval.

	1	2	3	4	5	
Very difficult	◯	◯	◯	◯	◯	Extremely easy

2. **How easy was to create a new project?**
   Mark only one oval.

	1	2	3	4	5	
Very difficuly	◯	◯	◯	◯	◯	Extremely easy

3. **How easy was to load the original files?**
   Mark only one oval.

	1	2	3	4	5	
Very difficult	◯	◯	◯	◯	◯	Extremely easy

4. **How easy was to locate the controls to start translating the document after loading it?**
   Mark only one oval.

1	2	3	4	5
◯	◯	◯	◯	◯

5. **How easy was to save your progress?**
Mark only one oval.

	1	2	3	4	5	
Very difficult	◯	◯	◯	◯	◯	Extremely easy

6. **How easy was to identify the new matches?**
Mark only one oval.

	1	2	3	4	5	
Very difficult	◯	◯	◯	◯	◯	Extremely easy

7. **How easy was to identify fuzzy matches and total matches?**
Mark only one oval.

	1	2	3	4	5	
Very difficult	◯	◯	◯	◯	◯	Extremely easy

8. **How easy was to use the terminology support (if any) while translating?**
Mark only one oval.

	1	2	3	4	5	
Very difficult	◯	◯	◯	◯	◯	Extremely easy

9. **How easy was to revise and proofread the document from within the software?**
Mark only one oval.

	1	2	3	4	5	
Very difficult	◯	◯	◯	◯	◯	Extremely easy

10. **How easy was to navigate through the controls to add, remove or modify the format of the document?**
Mark only one oval.

	1	2	3	4	5	
Very difficult	◯	◯	◯	◯	◯	Extremely easy

11. **How easy was to generate the final document?**
Mark only one oval.

	1	2	3	4	5	
Very difficult	◯	◯	◯	◯	◯	Extremely easy

12. **How easy was to work with text documents?**
Mark only one oval.

	1	2	3	4	5	
Very difficult	◯	◯	◯	◯	◯	Extremely easy

13. **How easy was to work with slideshows?**
Mark only one oval.

	1	2	3	4	5	
Very difficult	◯	◯	◯	◯	◯	Extremely easy

14. **How easy was to work with webpages?**
Mark only one oval.

	1	2	3	4	5	
Very difficult	◯	◯	◯	◯	◯	Extremely easy

Survey on CAT tools - Google translator toolkit

On a scale from 1 - 5 (difficulty) answer the following questions related to the ease of use of Google translator toolkit software

1. **How easy was to find the options that allowed you to create a new project?**
   Mark only one oval.

	1	2	3	4	5	
Very difficult	◯	◯	◯	◯	◯	Extremely easy

2. **How easy was to create a new project?**
   Mark only one oval.

	1	2	3	4	5	
Very difficuly	◯	◯	◯	◯	◯	Extremely easy

3. **How easy was to load the original files?**
   Mark only one oval.

	1	2	3	4	5	
Very difficult	◯	◯	◯	◯	◯	Extremely easy

4. **How easy was to locate the controls to start translating the document after loading it?**
   Mark only one oval.

1	2	3	4	5
◯	◯	◯	◯	◯

5. **How easy was to save your progress?**
Mark only one oval.

	1	2	3	4	5	
Very difficult	◯	◯	◯	◯	◯	Extremely easy

6. **How easy was to identify the new matches?**
Mark only one oval.

	1	2	3	4	5	
Very difficult	◯	◯	◯	◯	◯	Extremely easy

7. **How easy was to identify fuzzy matches and total matches?**
Mark only one oval.

	1	2	3	4	5	
Very difficult	◯	◯	◯	◯	◯	Extremely easy

8. **How easy was to use the terminology support (if any) while translating?**
Mark only one oval.

	1	2	3	4	5	
Very difficult	◯	◯	◯	◯	◯	Extremely easy

9. **How easy was to revise and proofread the document from within the software?**
Mark only one oval.

	1	2	3	4	5	
Very difficult	◯	◯	◯	◯	◯	Extremely easy

10. **How easy was to navigate through the controls to add, remove or modify the format of the document?**
Mark only one oval.

	1	2	3	4	5	
Very difficult	◯	◯	◯	◯	◯	Extremely easy

11. **How easy was to generate the final document?**
Mark only one oval.

	1	2	3	4	5	
Very difficult	◯	◯	◯	◯	◯	Extremely easy

12. **How easy was to work with text documents?**
Mark only one oval.

	1	2	3	4	5	
Very difficult	◯	◯	◯	◯	◯	Extremely easy

13. **How easy was to work with slideshows?**
Mark only one oval.

	1	2	3	4	5	
Very difficult	◯	◯	◯	◯	◯	Extremely easy

14. **How easy was to work with webpages?**
Mark only one oval.

	1	2	3	4	5	
Very difficult	◯	◯	◯	◯	◯	Extremely easy

OmegaT issues and difficulties

Please provide information related to issues and difficulties found upon wokrning with OmegaT

Observed issues

1. **What Problems did you perceive upon using the tool for the first time?**

2. **What difficulties did you perceive related to the access and management of terminology within the program?**

3. **What difficulties related to the access and management of the translation memory module did you perceive?**

4. **What difficulties did you perceive upon using the program under time constraints?**

5. **What difficulties did you perceive upon generating the translated file or document?**

6. **What difficulties upon using the project management module (if present)**

7. **Did you perceive any difficulties in regards to the interoperability between the software and the open source or proprietary formats of the documents? If so, what were they?**

.....

.....

.....

.....

.....

Observation notes and checklist format:

http://www.evernote.com/l/AGfJu59UhipEJ4aLZHSAFhnYtnm7lD9Kdhs/

Observation notes and checklist format - thesis

Notebook: !NBOX
Created: 17/01/2015 11:00 a.m. **Updated:** 04/03/2016 11:17 a.m.
Author: Emerson Mario Ovidio Sanchez Ruiz

- [ ] The student completed the translation of the (webpage, document, slideshow)
- [ ] The student created the project
- [ ] The student saved the project
- [ ] The student generated the final documents
- [ ] The student accessed the translation memory
- [ ] The student used the glossary if any (terminology database)

Notes:

Instruments

1. Timekeeping matrix
2. Mistakes found matrix

https://docs.google.com/spreadsheets/d/189JWccDdS2N_--6jyvsb953nkDFacJRCPbH1U_IU1-M/edit?usp=sharing

Mistakes found upon reviewing students work																		
document	Sample Text Document (.doc) 110 words (sample)						Sample Presentation (.ppt) 140 words (answering machine)						Sample web page (no format)(.htm) 41 words (01 intro)					
Tool	OMEGAT			Google Translator toolkit			OMEGAT			Google Translator toolkit			OMEGAT			Google Translator toolkit		
Types of mistakes	Grammatical	Punctuation	Syntax	Grammatical	Punctuation	Syntax	Grammatical	Punctuation	Syntax	Grammatical	Punctuation	Syntax	Grammatical	Punctuation	Syntax	Grammatical	Punctuation	Syntax
1	**2**	**1**	**0**	**2**	**2**	**1**	**1**	**1**	**1**	**0**	**1**	**1**	**0**	**0**	**0**	**0**	**0**	**0**
2	**1**	**3**	**2**	**0**	**0**	**0**	**0**	**2**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**1**	**0**
3	**0**	**1**	**0**	**2**	**2**	**1**	**1**	**0**	**1**	**1**	**1**	**1**	**0**	**0**	**1**	**0**	**0**	**1**
4	**2**	**0**	**0**	**1**	**0**	**1**	**1**	**1**	**2**	**1**	**1**	**0**	**1**	**0**	**0**	**1**	**0**	**0**
5	**0**	**2**	**2**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**
6	**0**	**0**	**0**	**1**	**0**	**0**	**1**	**0**	**1**	**1**	**0**	**0**	**0**	**0**	**0**	**0**	**1**	**0**
7	**0**	**1**	**1**	**0**	**2**	**1**	**0**	**2**	**2**	**0**	**0**	**1**	**0**	**0**	**1**	**0**	**0**	**1**
8	**2**	**2**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**1**	**0**	**0**	**0**	**0**	**0**	**0**	**0**
9	**0**	**1**	**2**	**0**	**2**	**1**	**0**	**0**	**2**	**0**	**0**	**0**	**0**	**0**	**1**	**0**	**0**	**1**
10	**1**	**1**	**0**	**1**	**0**	**1**	**1**	**0**	**0**	**1**	**1**	**0**	**1**	**0**	**0**	**0**	**0**	**0**
11	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**1**	**0**	**0**	**0**	**1**	**0**	**0**	**0**	**1**	**0**	**1**
12	**0**	**1**	**0**	**0**	**0**	**0**	**0**	**0**	**1**	**0**	**0**	**0**	**1**	**0**	**1**	**1**	**1**	**1**
13	**0**	**0**	**0**	**1**	**0**	**1**	**1**	**0**	**0**	**1**	**1**	**0**	**1**	**1**	**0**	**0**	**0**	**0**
14	**1**	**1**	**0**	**0**	**1**	**1**	**0**	**1**	**0**	**0**	**0**	**0**	**0**	**0**	**1**	**0**	**1**	**0**
15	**0**	**0**	**1**	**0**	**0**	**0**	**1**	**0**	**1**	**0**	**0**	**0**	**0**	**1**	**0**	**1**	**0**	**0**
16	**0**	**0**	**1**	**0**	**1**	**1**	**0**	**0**	**0**	**1**	**0**	**0**	**1**	**0**	**1**	**0**	**1**	**0**
17	**0**	**0**	**1**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**
	0.53	**0.82**	**0.59**	**0.47**	**0.59**	**0.53**	**0.41**	**0.47**	**0.65**	**0.35**	**0.35**	**0.24**	**0.29**	**0.12**	**0.35**	**0.24**	**0.29**	**0.29**
	9	14	10	8	10	9	7	8	11	6	6	4	5	2	6	4	5	5

	Sample Text Document (.doc) 110 words (sample)						Sample Presentation (.ppt) 140 words (answe		
Participant	OMEGAT			Google Translator Toolkit			**Participant**	OMEGAT	
	Active translation time	Time dedicated to revision	Time taken to finalize	Active translation time	Time dedicated to revision	Time taken to finalize		Active translation time	Time dedicated to revision
1	**0:15:04**	**0:04:00**	**0:02:00**	**0:12:00**	**0:04:00**	**0:04:00**	**1**	**0:33:00**	**0:04:00**
2	**0:12:00**	**0:02:00**	**0:02:00**	**0:07:00**	**0:04:00**	**0:04:00**	**2**	**0:11:00**	**0:02:00**
3	**0:19:21**	**0:02:00**	**0:02:00**	**0:09:00**	**0:09:14**	**0:04:00**	**3**	**0:14:00**	**0:02:00**
4	**0:17:40**	**0:02:00**	**0:02:00**	**0:10:00**	**0:04:00**	**0:04:00**	**4**	**0:40:00**	**0:07**
5	**0:18:25**	**0:02:00**	**0:02:00**	**0:05:00**	**0:09:14**	**0:04:00**	**5**	**0:32:00**	**0:02:00**
6	**0:10:03**	**0:02:00**	**0:02:00**	**0:02:00**	**0:04:00**	**0:12:00**	**6**	**0:23:00**	**0:02:00**
7	**0:13:10**	**0:02:00**	**0:02:00**	**0:07:00**	**0:10:00**	**0:04:00**	**7**	**0:37:00**	**0:03:00**
8	**0:10:12**	**0:03:00**	**0:02:00**	**0:07:00**	**0:09:14**	**0:04:00**	**8**	**0:29:00**	**0:03:00**
9	**0:32:45**	**0:02:00**	**0:01:00**	**0:07:00**	**0:04:00**	**0:12:00**	**9**	**0:17:00**	**0:02:00**
10	**0:09:14**	**0:02:00**	**0:02:00**	**0:07:00**	**0:04:00**	**0:04:00**	**10**	**0:15:00**	**0:07:00**
11	**0:15:00**	**0:02:00**	**0:02:00**	**0:10:00**	**0:12:00**	**0:04:00**	**11**	**0:10:00**	**0:02:00**
12	**0:30:02**	**0:02:00**	**0:02:00**	**0:07:00**	**0:04:00**	**0:04:00**	**12**	**0:25:00**	**0:08:00**
13	**0:18:00**	**0:04:00**	**0:02:00**	**0:07:00**	**0:04:00**	**0:04:00**	**13**	**0:06:00**	**0:04:00**
14	**0:12:00**	**0:02:00**	**0:01:00**	**0:07:00**	**0:12:00**	**0:04:00**	**14**	**0:30:00**	**0:08:00**
15	**0:12:40**	**0:04:00**	**0:02:00**	**0:07:00**	**0:12:00**	**0:12:00**	**15**	**0:10:00**	**0:04:00**
16	**0:15:00**	**0:02:00**	**0:02:00**	**0:07:00**	**0:04:00**	**0:03:00**	**16**	**0:05:00**	**0:05:00**
17	**0:12:10**	**0:02:00**	**0:02:00**	**0:07:00**	**0:04:00**	**0:04:00**	**17**	**0:08:00**	**0:02:00**
	0:16:03	0:02:25	0:01:53	0:07:21	0:06:41	0:05:21		0:20:18	0:03:56

				Sample web page (no format)(.htm) 41 words (01 intro)					
	Google Translator Toolkit			**Participant**	OMEGAT			Google Translator Toolkit	
Time taken to finalize	Active translation time	Time dedicated to revision	Time taken to finalize		Active translation time	Time dedicated to revision	Time taken to finalize	Active translation time	Time dedicated to revision
0:03:00	**0:08:00**	**0:10:03**	**0:08:00**	**1**	**0:08:00**	**0:08:00**	**0:08:00**	**0:08:00**	**0:08:00**
0:05:00	**0:09:14**	**0:08:00**	**0:08:00**	**2**	**0:08:00**	**0:08:00**	**0:08:00**	**0:08:00**	**0:08:00**
0:05:00	**0:10:12**	**0:10:03**	**0:09:14**	**3**		**0:07:00**	**0:08:00**	**0:08:00**	**0:08:00**
0:05:00	**0:08:00**	**0:07:00**	**0:08:00**	**4**	**0:10:12**	**0:08:00**	**0:08:00**	**0:08:00**	**0:08:00**
0:05:00	**0:10:12**	**0:08:00**	**0:07:00**	**5**	**0:08:00**	**0:08:00**	**0:04:00**	**0:08:00**	**0:08:00**
0:03:00	**0:09:14**	**0:10:03**	**0:09:14**	**6**	**0:08:00**	**0:08:00**	**0:08:00**	**0:08:00**	**0:08:00**
0:03:00	**0:08:00**	**0:07:00**	**0:09:14**	**7**	**0:08:00**	**0:08:00**	**0:08:00**	**0:08:00**	**0:08:00**
0:05:00	**0:09:14**	**0:08:00**	**0:07:00**	**8**	**0:10:12**	**0:08:00**	**0:08:00**		**0:08:00**
0:01:00	**0:12:00**	**0:08:00**	**0:04:00**	**9**	**0:08:00**	**0:08:00**	**0:08:00**	**0:08:00**	**0:08:00**
0:04:00	**0:10:12**	**0:08:00**	**0:08:00**	**10**	**0:09:14**	**0:07:00**	**0:04:00**	**0:08:00**	**0:08:00**
0:04:00	**0:08:00**	**0:07:00**	**0:07:00**	**11**	**0:10:12**	**0:08:00**	**0:08:00**	**0:10:12**	**0:07:00**
0:05:00	**0:08:00**	**0:08:00**	**0:12:00**	**12**	**0:08:00**	**0:07:00**	**0:08:00**	**0:08:00**	**0:08:00**
0:04:00	**0:12:00**	**0:07:00**	**0:08:00**	**13**	**0:08:00**	**0:08:00**	**0:08:00**	**0:08:00**	**0:07:00**
0:05:00	**0:08:00**	**0:08:00**	**0:07:00**	**14**	**0:09:14**	**0:08:00**	**0:08:00**	**0:08:00**	**0:08:00**
0:04:00	**0:10:12**	**0:10:03**	**0:08:00**	**15**	**0:08:00**	**0:08:00**	**0:08:00**		**0:08:00**
0:04:00	**0:08:00**	**0:08:00**	**0:07:00**	**16**	**0:10:12**	**0:08:00**	**0:08:00**	**0:08:00**	**0:08:00**
0:05:00	**0:08:00**	**0:08:00**	**0:08:00**	**17**	**0:08:00**	**0:08:00**	**0:08:00**	**0:08:00**	**0:08:00**
0:04:07	0:09:12	0:08:15	0:07:55		0:08:42	0:07:49	0:07:32	0:08:09	0:07:53

Time taken to finalize
0:08:00
0:07:00
0:07:00
0:07:00
0:07:00
0:08:00
0:08:00
0:08:00
0:07:00
0:07:00
0:07:00
0:07:00
0:07:00
0:07:00
0:08:00
0:08:00
0:07:00
0:08:00
0:07:25