Translators in the Loop: Observing and Analyzing the Translator Experience with Multimodal Interfaces for Interactive Translation Dictation Environment Design

Julián Zapata Rojas

Thesis submitted to the Faculty of Graduate and Postdoctoral Studies
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in Translation Studies

Under the supervision of
Dr. Jean Quirion

School of Translation and Interpretation
Faculty of Arts
University of Ottawa

Abstract

This thesis explores interactive translation dictation (ITD), a translation technique that involves interaction with multimodal interfaces equipped with voice recognition (VR) technology throughout the entire translation process. Its main objective is to provide a solid theoretical background and an analysis of empirical qualitative and quantitative data that demonstrate ITD's advantages and challenges, with a view to integrating this technique into the translation profession.

Many empirical studies in human-computer interaction have strived to demonstrate the efficiency of voice input versus keyboard input. Although it was implicit in the earliest works that voice input was expected to completely replace—rather than complement—text-input devices, it was soon proposed that VR often performed better in combination with other input modes. This study introduces multimodal interaction to translation, taking advantage of the unparallelled robustness of commercially available voice-and-touch-enabled multimodal interfaces such as touch-screen computers and tablets. To that end, an experiment was carried out with 14 professional English-to-French translators, who performed a translation task either with the physical prototype of an ITD environment, or with a traditional keyboard-and-mouse environment. The hypothesis was that the prototypical environment would consistently provide translators with a better translator experience (TX) than the traditional environment, considering the translation process as a whole. The notion of TX as introduced in this work is defined as a translator's perceptions of and responses to the use or anticipated use of a product, system or service.

Both quantitative and qualitative data were collected using different methods, such as video and screen recording, input logging and semi-structured interviews. The combined analysis of objective and subjective usability measures suggests a better TX with the experimental environment versus the keyboard-and-mouse workstation, but significant challenges still need to be overcome for ITD to be fully integrated into the profession. Thus, this doctoral study provides a basis for better-grounded research in translator-computer interaction and translator-information interaction and, more specifically, for the design and development of an ITD environment, which is expected to support professional translators' cognitive functions, performance and well-being. Lastly, this research aims to demonstrate that translation studies research, and translation technology in particular, needs to be more considerate of the translator, the TX, and the changing realities of the interaction between humans, computers and information in the twenty-first century.

Résumé

La présente thèse porte sur la traduction dictée interactive (TDI), technique de traduction en interaction avec des interfaces multimodales équipées de technologie de reconnaissance vocale (RV) tout au long du processus traductif. L'objectif principal est de présenter une base théorique approfondie et une analyse de données empiriques qualitatives et quantitatives rendant compte des avantages et des défis de la TDI en vue d'intégrer cette dernière à la profession de traducteur.

Bon nombre d'études empiriques en interaction personne-machine ont tenté de démontrer l'efficacité de l'interaction vocale vis-à-vis de l'interaction au moyen du clavier. Bien qu'il fût implicite dans les premiers travaux que la voix devait remplacer, plutôt que compléter, les dispositifs de saisie textuelle, on n'a pas tardé à observer que la RV était souvent plus performante en combinaison avec d'autres modes d'interaction. Notre étude intègre l'interaction multimodale à la traduction, tenant compte de la robustesse sans précédent des interfaces multimodales commerciales telles que les ordinateurs à écran tactile et les tablettes. À cette fin, une expérience a été effectuée auprès de 14 traducteurs professionnels qui traduisent de l'anglais au français. Les participants ont effectué une traduction en interaction soit avec la maquette d'un environnement de TDI, soit avec un poste de travail conventionnel à clavier et à souris. L'hypothèse formulée est que l'environnement-prototype de TDI est le gage d'une meilleure expérience traducteur (TX) que le poste de travail conventionnel, tout au long du processus traductif. La notion de TX, introduite dans cette recherche, est définie comme l'ensemble des perceptions et des réponses d'un

traducteur qui résultent de l'utilisation réelle ou anticipée d'un produit, un système ou un service.

Des données quantitatives comme qualitatives ont été récoltées au moyen de méthodes telles que l'enregistrement vidéo, des activités à l'écran et des données d'entrée, ainsi que les entrevues semi-dirigées. L'analyse combinée des mesures de convivialité tant objectives que subjectives suggère une meilleure TX avec l'environnement expérimental vis-à-vis du poste de travail à clavier et à souris. Cela dit, d'importants défis restent à surmonter pour que la TDI soit enfin intégrée de façon optimale à la profession. Ainsi, la présente étude doctorale fournit la base pour des recherches plus approfondies dans les domaines de l'interaction traducteur-machine et de l'interaction traducteur-information et, plus précisément, pour la conception et le développement d'un environnement de TDI; un tel système devrait améliorer les fonctions cognitives, la performance et le bienêtre des traducteurs professionnels. Enfin, la présente recherche servira d'exemple pour d'autres études en traductologie, et en traductique en particulier, lesquelles devront prendre en compte à la fois les traducteurs, la TX et les réalités changeantes des interactions des humains avec les machines et l'information au XXI^e^ siècle.

Acknowledgments

Foremost, I want to express a deep feeling of gratitude to Jean Quirion, who for many years has been more than an exceptional research supervisor. His involvement in my life has gone beyond the academic level; he has been a point of reference also at the professional and personal levels. He has managed to become one of a handful of true friends, which is way more than what one can expect from an advisor in this winding road of pursuing a doctorate. His professionalism and intellectual rigour are only overshadowed by his quality as a human being. His guidance and support on many aspects of my life have shaped this project and will follow me wherever it is I will go with my PhD in hand*. Merci, de tout mon cœur, Jean.*

To the thesis committee: Professor Heather O'Brien, from the University of British Columbia, and professors Elizabeth Marshman, Lynne Bowker and Marc Charron from the University of Ottawa, for their thorough examination of this work and their invaluable advice for future work. This thesis is only the beginning of the journey!

To Professors Robert Biddle and Avi Parush from Carleton University, for their involvement and guidance when consolidating my ideas and putting together a research project proposal. A special thanks also to Professor Pierre-Jérôme Bergeron from the University of Ottawa for willingly accepting to offer me invaluable advice on statistical considerations when designing my experiments.

My gratitude is also extended to all anonymous participants, and the anonymous organization they work for, for their vital involvement in the experiments. Their voluntary participation was invaluable for the project, which focuses on the human factors of technology tools. There is no better way to begin thinking about how to improve existing tools and designing new ones than when having human translators in the loop.

I also want to gratefully acknowledge the Language Technologies Research Centre, located in Gatineau, Canada, for being an essential part of this project by providing me with an office space and high-speed wireless Internet access at their facilities, free of charge, to perform the experiments.

Special thanks must go, too, to all my colleagues and acquaintances who also shaped this project by taking part in formal and informal experiments, by listening to my ideas and giving me feedback, by reading and proofreading portions of my writing, and by keeping me motivated.

I am also deeply grateful to that handful of closer friends who, for many years, have stood by me and walked next to me throughout the sinuous doctoral path, for being a motivation to strive to be better, for teaching me to never quit, for being there to support me in every way possible; simply for believing in me.

Last but not least, to my parents, my brother and my sister, who have been sources of inspiration and encouragement, *¡gracias!* Thank you for the very many sacrifices you have made for me. Thank you for your unconditional support, for your patience and praise. For being an essential part of who I am today. For believing in me.

The realization of this project could not have been possible without the generous financial support from the *Fonds québécois de la recherche – société et culture (FQRSC)* under grant 31791, and from the University of Ottawa.

Table of Contents

List of Abbreviations

ASR	Automatic speech recognition
BVR	Bilingual voice recognition
CAT	Computer-aided translation
CTDD	Conventional translation dictation with dictaphone
GUI	Graphical user interface
HCI	Human-computer interaction
HII	Human-information interaction
HTR	Handwritten text recognition
IB	Information behaviour
ICTs	Information and communication technologies
IR	Information retrieval
ITD	Interactive translation dictation
ITP	Interactive translation prediction
KM	Keyboard and mouse
LTs	Language technologies
MI	Multimodal interface
MT	Machine translation
NLP	Natural language processing
PC	Personal computer
SA	Speaker-adapted
SD	Speaker-dependent
SI	Speaker-independent
SiT	Sight translation
SR	Speech recognition

TAPs	Think-aloud protocols
TCI	Translator-computer interaction
TD	Translation dictation
TDVR	Translation dictation with voice recognition
TII	Translator-information interaction
TM	Translation memory
TMS	Translation memory system
TPR	Translation process research
TS	Translation studies
TT	Typed translation
TX	Translator experience
UI	User interface
UX	User experience
VR	Voice recognition
VTS	Voice, touch and stylus
WIMP	Windows, icons, menus and pointer

List of Figures

List of Tables

Chapter 1 — Introduction

"The problem is never how to get new, innovative thoughts into your mind, but to get old ones out. Every mind is filled with archaic furniture. Clean out a corner of your mind and creativity will instantly fill in." —Dee Hock

This thesis was written multimodally; no physical keyboard or mouse was used at any time. Instead, a combination of voice, touch, stylus and software keyboard input was used, on various computational platforms: a reclining touch-screen computer, a tablet and a smartphone. This was done either from a home office (sitting down or standing up) or from various Wi-Fi-connected venues such as coffee shops, hotels, airports, airplanes, and university study rooms and parks. The document was permanently accessible via a password-protected cloud-based drive and automatically synchronized when changes were made and the document was saved, on any of the devices. My doctoral study explores a similar multimodal configuration for the translation profession.

More specifically, this thesis explores interactive translation dictation (ITD). The main objective is to provide a solid theoretical background and an analysis of empirical qualitative and quantitative data that demonstrate ITD's current advantages and challenges, with a view to integrating it into the translation profession. Thus, the study makes an original contribution to the advancement of knowledge in the field of translation studies (TS) and, specifically, in translation technology.

ITD –a notion first introduced in my Master's thesis (Zapata Rojas, 2012)– is defined in the present work as a translation technique that involves interaction with multimodal interfaces (MIs) equipped with voice recognition (VR)[1] technology, throughout the entire translation process, namely during preparation, production and revision.

VR technology gives a machine the ability to recognize and process human voice and speech. Nowadays, although far from perfect, VR systems are widely used for dictation (i.e., for the transcription of speech by a single specific speaker) (Jurafsky & Martin, 2009, p. 285) in several major languages and in certain specific domains, and are also adapted to and used in a variety of daily-life situations (e.g., automated phone customer service and tech support; mobile phone activation and personal assistant interaction; in-vehicle navigation system interaction; live subtitling; search engine queries; etc.) (Haton, Cerisara, Fohr, Laprie, & Smaili, 2006; Jurafsky & Martin, 2009; Llisterri, 2009a, 2009b; Mariani, 2009; Parush, 2005; Romero-Fresco, 2011; Schalkwyk et al., 2010).

There are three types of VR systems with respect to speakers: speaker-independent (SI), speaker-dependent (SD) and speaker-adapted (SA). To build SI systems, acoustic data collected from many speakers across age, gender and dialect are used to train acoustic models, as well as speaker normalization techniques (see Uebel and Woodland (1999)). Thus, virtually any person can speak into an SI system. SD systems

[1] In some of the literature, the terms *VR*, *speech recognition (SR)* and *automatic speech recognition* (*ASR)* are used interchangeably to refer to the same notion. The same applies to *voice input* and *speech input*. I ask the readers to keep that in mind as they read the remainder of this thesis.

are built similarly to SI systems, but using only training data from a single speaker who will also be the sole user of the system. Lastly, SA systems constitute a middle road. The idea is to adapt an SI system to a specific user using only some speaker-specific data. At a practival level, speaker-adaptation happens in a supervised fashion. The user reads aloud a number of sentences to "train" the system and optimize its accuracy (see Woodland (2001) for a review of adaptation techniques). Most of commercially available VR systems are meant to be installed on personal computers (PCs), but recent developments in computing have made it possible to access any of these systems via a remote server or the cloud,[2] either from PCs or from mobile devices such as tablets and smartphones.

Many challenges have been associated with VR technology, mostly from the computational perspective. Back in the 1990s, Shneiderman (1998, p. 328) argued that VR is "the bicycle of user interface [UI] design: it is great fun to use and has an important role, but it can carry only a light load". Although significant advancements have taken place in computing since the 1990s, VR systems still require a great amount of computational power, i.e. the user must have a powerful device in order for the VR system to work properly. Likewise, training a VR system is a computationally expensive process and training commercial systems can only take

[2] The U.S. National Institute of Standards and Technology defines *cloud computing* as "a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g. networks, servers, storage, applications and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction" (Mell & Grance, 2011). Kurzweil (2013, p. 116) provides examples of cloud-based applications humans use on an daily basis: "When you use a search engine, recognize speech from your phone, consult a virtual assistant [...], or use your phone to translate a sign into another language, the intelligence is not in the device itself but in the cloud."

place on servers or clusters. That being said, as stated above, trained models can be embedded in a PC or can alternatively be used from a server accessed through the cloud.

The trade-off between embedded vs. cloud is one of computation vs. latency/connectivity. The computation required in speech decoding (i.e. the actual recognition) is also computationally expensive. This is not a problem for PCs plugged directly to a power source, but for laptops, tablets and smartphones, if the VR system is embedded in the device, decoding drains the battery and consumes memory to such an extent that VR may not be a practical tool. Work on reducing memory usage and still achieving acceptable VR accuracy has been conducted (see for instance Lei et al. (2013)). If speech is streamed to a server, decoded and the output returned via the cloud or intranet, electrical and computational power is abundant. However, the device must have web access fast enough to stream sound to the server and receive text. But latency (i.e. the time interval between the input and the output) in VR confuses users, who will stop dictating, repeat words, restart or speak slower than their natural speech rate. The problem of battery lifetime can be alleviated by users of a desktop PC, but this means that users are chained to the workstation, which is not appropriate for cases where mobility is desirable.

For over three decades, many studies in UI design and development have strived to demonstrate the efficiency of voice input versus keyboard-and-mouse (KM) input, and to overcome the technical challenges of stand-alone VR systems such as the ones described above. Although it was implicit in the earliest works that VR technology was expected to completely replace—rather than enhance or complement—text-input devices, it was soon proposed that voice input often performed better in combination with other input modes (Pausch & Leatherby, 1991). Jurafsky and Martin (2009, p. 285) explain:

> While many tasks are better solved with visual or pointing interfaces, speech has the potential to be a better interface than the keyboard for tasks for which full natural language communication is useful or for which keyboards are not appropriate [...].[3] [Now,] [i]n some applications, a multimodal interface [(MI)] combining speech [and other input modes] can be more efficient than a graphical user interface [GUI] without speech (Cohen et al. 1998).

The combination of voice with other input modes is indeed often thought by many researchers as more useful and even preferable for many applications, both from the user and computational viewpoints. Thus, in recent years, multimodal interaction has increasingly become a hot topic in human-computer interaction (HCI) research.

Multimodal interaction is at the core of my doctoral project. What characterizes MIs is their ability to process two or more combined user input modes (such as voice, stylus, touch, gaze, and gestures) in parallel and in a coordinated manner (Girol, Lopez-Cozar, Callejas, Espejo, & Ábalos, 2012; Oviatt, 2012; Williamson, Crossan, & Brewster, 2011). Examples of commercially available MIs are tablets, smartphones

[3] Nonetheless, the primary mode of textual data entry in most domains remains the conventional (mechanical, or physical) keyboard (Lee et al., 2012), and other devices such as chord keyboards are greatly used in fields where dictation is common. Chord keyboards allow several keys to be pressed simultaneously to represent several characters or even entire words. Courtroom recorders, for example, regularly use chord keyboards to transcribe the full text of spoken arguments, reaching rates of up to 300 words per minute (Shneiderman, 1998, p. 307).

and other touch-and-voice-enabled devices such as touch-screen computers. This new type of computational platforms represents a major and innovative departure from traditional GUIs, or WIMP (windows, icons, menus and pointer) interfaces, characterized by the KM as main input modes (Oviatt, 2012, pp. 418–419).[4]

Oviatt (2012, p. 415) argues that "[MIs] have the potential to expand computing to more challenging applications, to be used by a broader spectrum of everyday people, and to accommodate more adverse usage conditions than in the past." Today, with increasingly available MIs, we are just beginning to move towards a new computing era.[5]

The major impetus for developing MIs has been the practical aspects of mobile use; thus most of commercial MIs have primarily been designed for this purpose. But the interest in multimodal interaction goes well beyond mobility, as Oviatt (2012, pp. 422–423) explains:

> Ultimately, [MIs] are just one part of the larger movement to establish richer communications interfaces, ones that can expand existing computational functionality and also improve support for human cognition and performance [...]. One major goal of such interfaces is to

[4] Most current computer-aided translation (CAT) tools are GUI-based. This being said, some CAT tools are already mobile-friendly (i.e., they are designed for easy use on mobile devices such as smartphones and tablets). Examples are Canada's term bank Termium Plus: http://www.btb.termiumplus.gc.ca/tpv2alpha/alpha-eng.html?lang=eng (Retrieved on 02-03-2016); and Kanjingo, a mobile app for post-editing (O'Brien et al., 2014). CAT will be discussed in more detail in Chapter 2.

[5] Harper et al. (2008, pp. 14–15) from Microsoft Research describe the "four computing eras": the 1960s, or the Mainframe Era, characterized by the use of one computer by several people. Then, the 1980s, or the Personal Computer Era, that is, one computer per person. Thirdly, the 2000s, or the Mobility Era, in which one person can own several computers. Lastly, the Ubiquitous Era (2020 and beyond), in which one person will have access to thousands of computers, everywhere.

> reduce cognitive load and improve communicative and ideational fluency.

As far as the translation sector is concerned, MIs, and particularly voice-enabled interfaces, offer, in my opinion, great potential for translation students and professionals, since they can be more efficient than a GUI without speech in tasks for which full natural language communication is useful, as Jurafsky and Martin, quoted above, explain. This is precisely the case in translation practice, which is nonetheless currently characterized by the extended use of KM GUIs (Garcia-Martinez et al., 2014).

This being said, over the past two decades, certain researchers, trainers and professionals have demonstrated an interest in using VR technology in translation (Benis, 2002; Brousseau et al., 1995; Ciobanu, 2014; Désilets, Stojanovic, Lapointe, Rose, & Reddy, 2008; Dragsted, Hansen, & Selsøe Sørensen, 2009; Dragsted, Mees, & Hansen, 2011; Dymetman et al., 1994; Gouadec, 2007; Mees, Dragsted, Hansen, & Jakobsen, 2013; Mesa-Lao, 2014; Reddy & Rose, 2010; Rodriguez, Reddy, & Rose, 2012; Romero-Fresco, 2012; Seaman, 2002; Stroman, 2002; Vidal, Casacuberta, Rodríguez, Civera, & Martínez Hinarejos, 2006). However, none of today's off-the-shelf stand-alone VR systems was developed specifically for translation purposes (Ciobanu, 2014; Gouadec, 2007). The technological limitations identified include the inability of off-the-shelf VR systems to work bilingually (that is, in two languages concurrently),[6] to carry out translation-specific tasks, and to work together with computer-aided translation (CAT) tools such as translation memory systems (TMSs).

This thesis is intended as a first step towards overcoming the challenges and limitations of VR technology for the translation sector by introducing multimodal interaction. It takes advantage of the unparallelled robustness of commercially available (voice-enabled) MIs. To that end, an experiment involving MIs and professional translators was carried out. A literature review and a discussion on various topics, including translation dictation (TD), CAT, bilingual voice recognition (BVR) and multimodal interaction, will also be provided in this thesis.

Let us now offer a brief overview of the two subfields that will be used as a theoretical framework for the present doctoral study: translator-computer interaction (TCI) and translator-information interaction (TII). A deeper discussion will be provided in Chapter 2.

1.1. Translator-Computer Interaction and Translator-Information Interaction: Two Emerging Subfields in Translation Studies

In our day, translation practice is both an information-interaction task and a computer-interaction task; therefore, it is vital to closely examine and apply the theories, research principles and methodologies of such fields as

[6] Other natural language processing (NLP) applications now have the ability to handle two languages in parallel within the same interface. An example of this is the Antidote 9 software by Druide Informatique, a full-fledged language tool in both English and French, which includes a set of dictionaries (collocations, synonyms, etc.), a spelling and grammar checker, and a set of interactive language guides. For more information, see http://www.antidote.info/antidote/for-english-speakers. Retrieved on 02-05-2016.

human-information interaction (HII) and HCI. These two fields have the potential to better inform all research aiming to provide translators with advanced tools that can effectively and efficiently[7] assist them in carrying out those information-interaction and computer-interaction tasks.

HII and HCI carry two complex and slippery notions difficult to understand, identify and define in an exact way: *information* and *computer,* respectively. Since these notions are dealt with in an array of disciplines and by hundreds of researchers, it is not surprising that both present various definitions and interpretations.

In regard to *information*, Fidel (2012, p. 7) explains that the concept

> has been a source for many theoretical discussions. Both the context in which it is used and the philosophical stance of the scholar who defines it shape the concept's construal. […] It is reasonable, then, to accept *information* as a concept with various types, where each type caters to a certain context.

In the context of my research, which stems from my interest in the design of a system that improves the interaction of human translators with information, I adopt the definition of *information* articulated by Fidel (*ibid*.) as a string of symbols (e.g. alphanumeric) that has meaning, is communicated, has an effect, and is used for

[7] In the field of HCI, *effectiveness* is the accuracy and completeness with which users achieve specific goals, and *efficiency* is the resources expended in relation to the accuracy and completeness with which users achieve goals (ISO, 1998).

decision making. A word, a term record, a dictionary definition, a newspaper article, or any piece of writing are thus examples of information.

As far as the notion of *computer* is concerned, it is also worth noting that as fundamental as it is, a broad and comprehensible definition is difficult to find. In the early stages of computing, around the 1940s, *computer* was actually a job description for a person who performed mathematical operations for large-scale projects. When the first machines capable of performing such operations emerged, the term *computer* was used to name those machines as well.[8] In a 1997 essay, Hayes provides an account of an e-mail discussion he had with several field experts about "What is a Computer?" (Hayes, 1997). Such an account demonstrates the complexity of the task of defining *computer*. For instance, some experts provide definitions as broad as "something that implements a function", or "a machine which performs computations", while several others have put forward the idea that the human brain is, too, a computer (see Hayes' essay for further discussion). While I acknowledge the complexity of this matter, I will use, for the purposes of my research, a straightforward definition by Microsoft: "Computers are machines that perform tasks or calculations according to a set of instructions [by the user], or programs."[9] In this thesis, desktop PCs, laptops, smartphones and tablets are examples of computers.

Let us now go back to the core notions of HII and HCI. HII is defined as the field of study that investigates the way humans interact with information (Fidel, 2012;

---

[8] I learned this information while visiting Harvard University, in Cambridge, Massachusetts, where Mark I (the first programmable computer in the United States) is displayed.

[9] Source: http://windows.microsoft.com/en-ca/windows/introduction-to-computers#1TC=windows-7. Retrieved on 23-06-2016.

Gershon, 1995). It constitutes an emerging discipline that encompasses several subfields, including HCI. Hewett et al. (1992, p. 5) define the latter as the "discipline concerned with the design, evaluation and implementation of interactive computing systems for human use and with the study of major phenomena surrounding them". HCI is closely related to the field of ergonomics (Hewett et al., 1992, p. 8; Vicente, 1997, 2004), also known as human factors. The International Ergonomics Association[10] defines the latter as

> the scientific discipline concerned with the understanding of interactions among humans and other elements of a system,[11] and the profession that applies theory, principles, data and methods to design in order to optimize human well-being and overall system performance.

According to Sáenz Zapata (2011, p. 155), given its anthropocentric approach and perspectives, ergonomics has become an increasingly important discipline in recent years not only in the workplace, but also in the realms of academia, science, and the day-to-day lives of humans.

In addition to ergonomics, HCI is informed by several other disciplines in the technologies and in the human and natural sciences, in particular, psychology, physiology, anthropology, engineering, industrial design and computer science. Their common goal is to "humanize" (Oviatt, 2012, p. 423) the technology in search of a

[10] Source: http://www.iea.cc/whats/. Retrieved on 27-11-2015.

[11] It is worth noting that a "system" here may not refer only to a machine or computer. As sociologists like Bruno Latour (1996, 1997) and Michel Callon (2006) put it, together with humans, machines are only part of a system or network, the elements of which are in constant interaction. For instance, a system in a school classroom includes students, teachers, their ideas, and technologies (tables, chairs, computers, etc.).

more natural interaction between humans and computers. Natural interaction leads to a better user experience (UX).

ISO (2010) defines UX as "[a] person's perceptions and responses that result from the use or anticipated use of a product, system or service". Harper et al. (2008, p. 56) explain:

> In recognition of the way computer use is changing, a number of [HCI] researchers and practitioners have begun studying the nature of the [UX] and how it unfolds over time. This has largely involved defining its subjective qualities, such as [how it feels] interacting with a device […]. Concepts such as pleasure, aesthetics, fun and flow, on the one hand, and boredom, annoyance and intrusiveness, on the other, have been used to describe the multifaceted nature of such 'felt' experiences.

In short, HCI is the search for computer applications that are *usable, useful* and *universal.* Usable, by "making the next generation of interfaces more appealing and the end users more successful" (Shneiderman, 2008, p. XX); useful, by making technologies that "will make it possible to deliver vital services [...][,] that more effectively harness each individual's capacity to contribute and participate in shaping the future" (*ibid.*); and universal, by making technologies that "enable people with diverse backgrounds to become more successful satisfied users: old & young, men & women, novice & expert, users with varied disabilities, and users with low motivation, low self-confidence, and poor reading skills" (*ibid*.).

Now, according to some researchers, HII constitutes a larger discipline, since it looks beyond computers; it focuses on the interaction between humans and the information in the universe, in all its complexity, regardless of the tools used to mediate or facilitate such interaction (Fidel, 2012; Gershon, 1995; Marchionini, 2008). In fact, HII emanated from the Information Science sub-fields of information behaviour (IB) and information retrieval (IR). Its importance is illustrated by the number of scientific papers using the term HII, its inclusion in university curricula, and the emergence of research centres devoted to its study (Marchionini, 2008, p. 173). (Further discussion on HII to be provided in Chapter 2).

Logical extensions of HII and HCI respectively, two notions will be at the core of the present work: translator-information interaction (TII) and translator-computer interaction (TCI). The notion of TII was introduced in Zapata (2016) as the field of study that investigates the interaction of human translators with (digital) information and information tools. This emerging notion complements that of TCI, introduced by O'Brien (2012): the study of the interaction between translators and computers. Now, although the notions of TCI and TII are fairly new, the interaction of human translators with machines and digital information is far from being a recent phenomenon (O'Brien, 2012, pp. 103–104):

> Already with the introduction of the electronic typewriter, with only two lines of memory, and the use of dictaphones, translation became a computer-interactive task. This was followed by the introduction of word-processing software [...][,] a development that would have required some translators to interact with a computer for the first time. Not long after the mass embracing of word processing, came the introduction of Translation Memory [TM] tools [and] terminology management programs, which are [...] not restricted to the storage of terms, but also store phrases and sometimes even sentences or larger chunks of text [...].

Hence, my doctoral study explores new avenues in HII and HCI from the perspective of translation in an effort to consolidate TII and TCI as pertinent subfields. The aim of the present work is to provide a solid theoretical background and empirical evidence of the potential that VR offers in combination with other recognition-based technologies, as well as wireless/mobile technologies and web-based tools and information, to the translation sector. Indeed, these emerging technologies and applications already provide professionals in other domains –and humans in general– with new, efficient, alternative ways of *interacting with computers* and of *accessing, using, and communicating information*. Why not in translation?

1.2. Objectives

The general objective of this doctoral thesis is to provide a solid source of theoretical and empirical evidence that can inform TII and TCI research and, consequently, the design and development of new tools that will better address the genuine needs of translators in the twenty-first century. Carl et al. (2011) argue that the

> [d]evelopment of translation tools could benefit from incorporating knowledge of human translation behaviour and translator styles [...]. Better cognitive models of translation have to be developed, since experiments to design and evaluate [such tools] are extremely labour intensive [...]. As Knight et al. (2007)[12] point out, "the combination of small usability studies and cognitive modelling [may help to] make an informed decision about critical aspects of a [UI]."

[12] Knight, A., Pyzark, G. & Green, C. (2007). *When Two Methods Are Better Than One: Combining User Study With Cognitive Modeling.* In *CHI'07 Proceedings*, pp. 1783-1788.

The specific objective of the study is to examine the possibility to translate, not with keyboard and mouse, but with VR in combination with other emerging input modes such as touch and stylus, on off-the-shelf MIs. The study should provide a basis for designing, developing and testing the working prototype of an ITD environment in the near future.

1.3. Research Question and Hypothesis

This research project is motivated by an underlying research question and a hypothesis, which stem from my preliminary investigations, observations and pilot experiments.

The research question which will be used as the basis of my study is formulated as follows:[13]

> *Does the physical prototype of an interactive translation dictation environment consistently provide translators with a better translator experience than a traditional keyboard-and-mouse environment, when preparing, producing and revising a translation?*

My hypothesis is formulated as follows:

> *The prototypical interactive translation dictation environment consistently provides translators with a better translator experience than the keyboard-and-mouse environment, considering the translation process as a whole.*

[13] A discussion of key notions in the research question and hypothesis is provided below.

1.3.1. A Note on Investigating Usability and Translator Experience

The notion of translator experience (TX) is introduced in this research. Based on ISO's definition of UX (2010), *TX* can be defined as a translator's perceptions of and responses to the use or anticipated use of a product, system or service.

According to Bargas-Avila and Hornbæk (2011), UX has been used as an umbrella term "for new ways of understanding and studying the quality-in-use[14] of interactive products". The authors explain that, commonly, the term UX has been used in the HCI literature to refer to the design and use of UIs, working thus as a synonym for usability. Nielsen (2012) defines usability as "a quality attribute that assesses how easy [UIs] are to use. [It] also refers to methods for improving ease-of-use during the design process."

Studying and improving the usability of interactive systems has long been a key element of HCI research. Although thoroughly measuring usability is considered a highly challenging task, it has been possible for researchers to operationalize the usability construct to find aspects of usability that can actually be measured and analyzed (Hornbæk, 2006). Traditionally, usability studies have reported on three aspects (*ibid.*):

1. Effectiveness (measures of the quality of or satisfaction with the outcomes of the interaction);
2. Efficiency (measures of the interaction process); and/or

[14] Bevan (1995) defines quality-in-use as "the extent to which a product satisfies stated and implied needs when used under stated conditions."

3. User satisfaction (measures of users' attitudes and experience).

Now, it has been proposed that, for HCI and HII studies to suceed, a better assessment of the correlations between the three aspects of usability described above needs to be done, and a better understanding of the correlations between objective and subjective measures needs to be achieved (*ibid*.). Hornbæk discusses how much most studies on usability actually fall short of a reliable account of the UX and of the "quality-in-use" of the interface or product investigated; only a partial view of usability is generally presented.[15]

Beyond usability, the main focus of the present-day UX movement is on the experience of interaction with interactive products, instead of the products themselves (Bargas-Avila & Hornbæk, 2011). Thus, another, most accurate use of the term UX is to designate a new paradigm in research that focuses on "non-instrumental needs and experience in a more complex sense" (*ibid*.).

In the present study, to confirm or falsify my hypothesis, I will examine both objective measures including expert assessment of the outcome, task times and usability patterns; and subjective measures such as users' perception of the outcome, subjectively experienced time and perception of task difficulty. I will also attempt to demonstrate how the inclusion of qualitative data, and a combined analysis of these and quantitative data may provide a better account of the actual interaction

[15] Based on a review of 180 studies published in HCI journals and proceedings, Hornbæk provides in his 2006 paper a critique of the way usability is commonly perceived, measured and assessed. A deeper discussion on the challenges of usability research will be provided in Chapter 5.

experience, highlighting thus a paradigm shift away from existing usability research, which is “too focused on task efficiency and work” (Bargas-Avila & Hornbæk, 2011).

1.3.2. A Note on Physical Prototyping

Prototyping is a vital part of the design and development process of any application or product. To arrive at usable UI designs, designers commonly build prototypes (i.e., approximations of a product). A prototype embodies a design hypothesis and enables a designer to test it (Hartmann et al., 2006). Given the scope and limitations of a doctoral project in TS (see section 1.6 below), I am using the physical prototyping technique. Applied early in the design process, physical prototyping utilizes and combines various existing tools and materials.

In sum, I am not intending to *build* a system as part of my doctorate, but to collect quantitative and qualitative data from translators using different off-the-shelf devices and software, in order to inform the design and the development of an eventual tool.

1.4. Methodology Overview

A brief overview of the methodology used in this doctoral study is provided in the present section. The experimental framework will be described in further detail in section 5.3.

In order to confirm or falsify my hypothesis, this doctoral project followed the path described below.

First, a thorough review of existing literature in HII, HCI and TCI was carried out, and a number of formal and informal pilot experiments were conducted, which were key elements in the design of the main experiment. A good part of this literature review and experiments was carried out as part of class projects and assignments both at the University of Ottawa (doctoral seminars) and at Carleton University, where I took two introductory graduate-level courses in HCI during the first session of my PhD.

The core of the thesis is the user study for ITD environment design: the observation and analysis of the TX with MIs. Considering the scope and limitations of a doctoral project (in the humanities) and the various challenges inherent in the research methods and tools commonly used in translation process research (TPR) (see section 5.2.1), the study includes active observation of users with video recording, screen recording, input logging and semi-structured interviews. It is also worth mentioning at this point that two subgroups of participants took part in the experiments: a test or experimental group (who translated a short text using MIs) and a control group (who translated the same text using traditional input modes, i.e., KM).

A physical prototype of an ITD environment was built. This prototype consisted of two commercial MIs, with which the experimental group interacted through – and only through – voice, touch and stylus in order to translate a short text from English into French: a 10" Android tablet and a 27" Windows touch-screen computer. The tablet was equipped with a cloud-based SI VR system integrated in a fast-typing virtual keyboard app (Swype, by Nuance). Participants in the test group used this device to perform searches online (e.g. by using Google Search, either manually using the Swype keyboard or using the voice search feature) and to access online CAT tools, such as terminology databases, dictionaries and corpora, while preparing and/or revising their translation. The touch-screen computer was equipped with a text-processing application (Microsoft Word 2013), with a special toolbar[16] installed (an add-on for Word developed by Terminotix Inc. specially to provide language professionals with direct access to informational resources from the MS Word interface); a state-of-the-art commercial SA VR system (Dragon NaturallySpeaking 13, by Nuance); and Windows 8's virtual (or software) keyboard. Participants used this interface to read/prepare the source text and to dictate and/or type (with the software keyboard) the target text and make corrections or changes to it. In addition, by means of voice shortcuts, subjects were able to perform searches on two informational resources directly from the source-document in Word using the Terminotix toolbar.[17]

As mentioned above, a control group was also recruited. Participants in the control group translated the same text using conventional input modes (i.e., KM on a desktop PC) and were instructed to translate the text "as they would normally do" at work.

Once the analysis phase began, all digital video, audio, screen recordings and input log data collected during the experiments were used to analyze the TX and to compare results. Video recordings helped in timing some individual tasks and made it possible to observe the participants' interaction with the different interfaces, tools and devices. Input logging data and screen recordings made it possible to have a closer

[16] I am referring to the Terminotix Toolbar for MS Word, developed and made available for free by Terminotix, Inc., a Canadian CAT tool provider. The Terminotix toolbar can be downloaded from the following address: http://www.terminotix.com/index.asp?content=brand&brand=13&lang=en.
[17] More details on the methodology will be provided in section 5.3. The experimental environment is illustrated in Figure 9 on page 107.

look at the interaction with the different input devices and the various applications and information sources used while completing the translation task, and to conduct a fairly detailed response time and error analysis. Lastly, audio recordings made it possible to transcribe and study the participants' answers to interview questions.

The experiment was intended to investigate whether the combined use of voice, touch and stylus (VTS) input puts forward a better TX, in every step of the translation process, than conventional input modes, and to examine where the speed bumps and frustrations in these interactions lay. Thus, in light of such an analysis, it should be possible later to build conceptual models (Parush, 2015) and formulate recommendations for the design of an ITD environment in order to build, in collaboration with a team of developers and a commercial partner, a functional prototype in the near future.

1.4.1. A Note on the Choice of Interaction Modes

Let us note that, although other modes of interaction are possible in translation, within the framework of this study, I focus particularly on VTS interaction. The interest in voice input was discussed earlier in this chapter. While VR allows the user to interact with the interface through voice, and touch may be used for pointing, the stylus may be used, in addition to pointing, to simulate the use of a pen on paper, either for writing or for drawing. Lastly, on multimodal VTS systems, it is even possible to alternate between VR, stylus input and touching directly on the screen with the fingers, increasing thus the multimodal possibilities for the user. Further discussion on multimodality will be offered in Chapter 4.

1.5. Ethics Approval

As per university and governmental regulations in Canada, due to the inclusion of human participants in this study, the project was submitted for approval to the University of Ottawa's Research Ethics Board in the first quarter of 2014. The project entitled *Integrating the Human Factor in Translation Tool Design: A Comparative Analysis of Various Translators' User Experience with Multimodal Interfaces* was granted approval on April 16, 2014. For full ethics approval notices, see appendix A on page 208.

1.6. Understanding the Project's Limitations

This study is carried out within the framework of a doctorate in TS. Therefore, a number of limitations apply, owing mostly to time constraints and human and financial resources.

1.6.1. Time and Financial Limitations

A doctoral thesis must be completed within a limited number of years; in the case of this project, governmental funding was available for a period of three years starting from the third academic session. I had no intention nor did I possess the resources to go beyond this limit. Financial and time limitations push a researcher to determine a manageable scope from the start, to set realistic goals and to limit the presentation and the analysis of data. In addition, for this project, it was necessary to take into consideration the time and effort needed to obtain an approval from the research ethics board, and to recruit human participants.

1.6.2. Participant Recruitment and Sampling

Human-centred research entails many challenges. Not only is it necessary to go through the long process of obtaining a research ethics approval, but also to recruit participants, which is not an easy task. Indeed, immediately after receiving the approval from the research ethics board in April 2014, a call for participation was sent out to a number of translation agencies and organizations in the Ottawa and Montreal regions (with the approval of their directors), and to individuals in my extended professional network. From nearly 200 translators, according to my calculations, who received the invitation, only a handful expressed an interest in taking part of the experiment, and only two managed to take some time off to perform the experiment, in late June 2014. The data collected from these two individuals was discarded from the thesis (since the experimental setup changed considerably later on, e.g. because newer versions of the VR software used were released a few weeks later).

The recruitment process needed to be revised. After approval from the research ethics board, I managed to obtain the involvement from a translation organization in the Ottawa Region, which had not been previously contacted. The organization's management agreed to distribute the recruitment e-mail among their employees. A few weeks later, I was able to recruit 14 participants, all professional translators, half of whom held management positions within the organization at the time of the experiment. The experiment took place over three weeks in October 2014.

The differences in the participants' profiles pose a major problem for performing statistical analyses in empirical research. The many variables (age, years of experience, type of position they hold) were counterbalanced to ensure, as much as possible, an equal distribution of participants into the two groups, that is, the control group and the experimental group. I acknowledge that the sample may not represent the entire population of translators accurately. Furthermore, 14 participants (seven in each group) may be considered a small sample for any statistical analyses to be valid. Therefore, the results of this study cannot be used in generalizations pertaining to the entire population of translators in every context and every language combination.

1.7. Contribution to the Advancement of Knowledge

This project will make an original contribution to the body of knowledge in TS. In recent years, there has been an increased recognition of the importance of technologies and the role they play in translation research, teaching and practice. More than ever before, translation researchers, trainers and professionals are aware of the need to improve existing tools and create new tools to cope with the evolution of technology and the ever-changing professional needs of translators.

On the one hand, current translation tools often lead users to frustration (LeBlanc 2013; Taravella 2013) and a full integration of those tools to the profession has not yet been achieved (Taravella & Villeneuve 2013). On the other hand, dictation has often been an ideal strategy in any writing activity, and has historically proven effective and efficient in translation (Dragsted & Hansen, 2009; Dragsted et al., 2009; Hétu, 2012; Laroque-Divirgilio, 1980, 1981; Mees et al., 2013).[18] Likewise, VR

[18] A more detailed discussion on dictation will be provided in Chapter 3.

technologies have increasingly achieved higher performance and maturity levels in several major languages such as English and French, among others, to the point of becoming appealing, efficient and ergonomic in certain professional domains and even, to a certain degree, in professional translation, since it "quite simply makes life much more comfortable because you are not physically tied to your keyboard or screen" (Benis, 2002, p. 27). In light of this, it seems more than appropriate to consider an optimal integration of VR technology to the translator's toolbox (i.e., the adaptation of this technology to the array of tasks that translators perform at work).

In our day, software designers and developers want to create efficient, user-friendly, visually appealing systems (Shneiderman & Plaisant, 2010). As far as the translation sector is concerned, there is a tangible need (Carl et al., 2011; Ehrensberger-Dow & O'Brien, 2015; O'Brien & Ehrensberger-Dow, 2013; O'Brien, 2012; Taravella & Villeneuve, 2011, 2013) to design and develop ergonomic and flexible interfaces that take the human factor into consideration,[19] and that are adapted to the translator's workflow, since any technology application that is too rigid impedes the work that it is meant to support (Karamanis, Luz, & Doherty, 2011, p. 49). Shneiderman (1998, p. 16) argues that

> [t]he enormous interest in human factors of interactive systems arises from the complementary recognition of how poorly designed many

[19] I invite the reader also to consult the final report of an interdisciplinary study under the name of ErgoTrans (Cognitive and Physical Ergonomics of Translation), carried out between 2013 and 2015 at the Zurich University of Applied Sciences with the support of the Swiss National Science Foundation. The report may be accessed online at: https://www.zhaw.ch/storage/linguistik/forschung/uebersetzungswissenschaft/ergotrans_final_report.pdf. Retrieved on 22-06-2016.

> current systems are and how genuinely developers desire to create elegant systems that serve the users effectively.

A typical HCI research project seeks to design or redesign a particular computing technology in order (1) to improve upon or enhance a given experience or (2) to create a different experience than before (Harper et al. 2008, 58):

> In both situations, initial research is conducted by learning more about people's current experiences […]. Ethnographic studies, logging of user interaction and surveys are commonly deployed. Based on the findings gathered, we begin to think about why, what, and how to design something better. To aid the process, usability and [UX] goals are identified and conceptual models developed. Prototypes are built, evaluated, and iterated, demonstrating whether the user goals have been met or whether the new [UX] is judged to be enjoyable, pleasurable or valuable by the target group.

Thus, user studies combining tool usage, translation processes and the TX are therefore more than necessary in translation research. As O'Brien (2012, pp. 116–117) puts it:

> "[TCI] would likely benefit from an increased focus on ethnographic-style, cognitive ergonomic studies of both translation tools and the translation process itself […]. More experimental studies of translator-tool interaction could be carried out using formal usability research methods such as screen-recording, eye tracking, and observation, the results of which could then be used by translation technology developers to improve the specifications of tools for the benefit of translators and, ultimately, the end users of those translations.

This study is as such a major contribution to the advancement of knowledge in the field of TS, and will have, without the slightest doubt, a positive impact on the translation sector and, most importantly, on human translators.

1.8. Outline of the Thesis

The remainder of this thesis is organized as follows:

Chapter 2 presents a general overview of TCI and TII at the era of information and communication technologies (ICTs) by describing the evolution of the interaction between translators, computers and information, and the evolution of reflections on this matter among the research community.

Chapter 3 provides a classification of different translation techniques including typed translation (TT) and TD, as well as a discussion on their advantages and current challenges.

Chapter 4 explores VR technology and multimodal interaction from the perspective of translation. It first explores bilingual VR (BVR). It discusses whether BVR ought to be considered in TCI and TII research, and particularly in ITD environment design and development. Subsequently, the chapter provides an in-depth discussion of multimodal interaction by presenting its evolution and a number of experiments that account for its advantages and challenges. Lastly, a close-up look at ITD is provided.

Chapter 5 describes the experimental framework of this study. It first discusses the challenges of measuring usability and presents the measures used in the experiment. It then offers a look at some data collection techniques commonly used in TPR, and a detailed description of the methodology used to collect empirical quantitative and qualitative data for this thesis.

Chapter 6 analyses and discusses the data collected in the experiment.

Lastly, Chapter 7, the concluding chapter, summarizes the reflections and findings of the thesis, and outlines avenues for future research.

Chapter 2 — Translation in the Era of Information and Communication Technologies: From Word-Processing to Interactive Translation Dictation

"The wonderful thing about modern technology is the amount of communication and information-sharing it facilitates. And the awful thing about modern technology is the amount of communication and information-sharing it facilitates."

—Mark McGuinness

This chapter explores translator-computer interaction (TCI) and translator-information interaction (TII) in the era of information and communication technologies (ICTs). First, it highlights the impact that ICTs have had on the translational landscape, leading to constant debates among translation studies (TS) researchers, and therefore to ongoing research efforts and technological innovations. Secondly, the chapter argues why the study of the interaction between translators and information is essential in current TS research.

2.1. Translator-Computer Interaction

It is difficult to think about translation today without thinking about computer tools and technologies. ICTs, and in particular language technologies (LTs), have been part of the translational landscape for many years now, and have inexorably modified the way translation is perceived, taught, practised and used. Bowker (2006) explains:

> As the volume of text to be translated increases, so does the pressure on translators to be able to work quickly and efficiently. Many translators are turning to technology in the hopes that it can help them increase their productivity. Whenever a new way of working is

> introduced, there will inevitably be effects on both the process and the product.

ICTs are defined as technological applications based on computing, microelectronics, telecommunications and multimedia, the combination and interconnection of which allow people to search, capture, process and transmit data of different natures (text, audio, image, video, etc.); to interact with each other and with machines; to access information; and to spread and share information (Touré, Mbangwana, & Sène, 2009, p. 35).

LTs are part of ICTs. LTs are defined in this thesis as applications that facilitate the active or passive use of one or more natural languages. Certain LTs are developed for the general public, while others are developed for language professionals (e.g. writers, translators, terminologists).

ICTs and LTs have become essential objects of study in TS research as humans (and therefore human translators) become more and more submerged in the digital world. Likewise, researchers, trainers, practitioners and other stakeholders need to adapt to the societal changes provoked by the introduction of new technologies.

Let us now discuss this matter in further detail.

2.1.1. The Importance of Language Technologies and Information and Communication Technologies in Translation Studies

Nowadays, LTs and ICTs in general play a crucial role not only in translation practice but also in TS research. Indeed, because translation is essentially a computer-interaction task (or historically, a tool-interaction task, ever since translation has been performed in written form; see Chapter 3), *translator* and *technology* are two elements that must be considered and analyzed together in the theorization of translation. Callon (2006, p. 267) explains that technological tools are endogenous factors in the socio-technical network (*réseau sociotechnique*); they become one with humans. As translators become acquainted and experienced with various technological tools, these tools are no longer exogenous factors; rather, they assume their role as endogenous factors in the system. Therefore, as Taravella and Villeneuve (2011) put it, “the theorization of translation must […] reposition the translator and the technologies in the same dimension, in the same production unit.”[20, 21]

According to Callon (2006, p. 268), technological artifacts are shaped by “external forces, the origin of which is placed within society: they may be, for instance, interests or ideologies or even social relations of domination and power."[22] Thus, changes in society lead to technological innovations, and vice versa. There is, argues Callon, an ongoing negotiation between humans and non-humans in order to conciliate the changing needs of each actor of the socio-technical network. This is why the role of technologies in TS research, and in the humanities ans social sciences in general, remains crucial (*ibid*., p. 272):

> The contribution of non-humans can no longer be ignored or minimized in the social sciences, since more and more investments are taking place in research and innovation, leading to an exponential

[20] *La théorisation [de la traduction] doit donc repositionner [le traducteur] et les technologies dans une même dimension, dans une même unité de production.*

[21] All translations into English from French-language sources are my own throughout this thesis.

[22] *Des forces extérieures, dont l’origine est placée dans la société : il peut s’agir par exemple d’intérêts ou d’idéologies ou encore de relations sociales de domination ou de pouvoir.*

> increase in the number of developments. [This being said,] acknowledging the active role of non-humans leads to several consequences.[23]

One of the many consequences of acknowledging the importance of technologies in the theorization of translation is the conflict of interpretations (i.e., different individuals may see different things from different viewpoints), leading to multiple opinions and reactions and, therefore, to long debates within the research community (Arnold, 2003; Bowker, McBride, & Marshman, 2008; Bowker, 2011; Church & Hovy, 1993; Garcia, 2009; LeBlanc, 2013; Macklovitch & Russell, 2000; Muegge, 2013; Pym, 2005, 2011; Taravella & Villeneuve, 2013). Some of these debates take place around how appropriate technologies are in general, and the impact they have on human performance and cognition; other debates focus on specific applications and their impact on the translation industry, on translated texts and on translators, as discussed in the following few subsections.

2.1.2. The Debate on the Pertinence of Technology

Attitudes towards technology are not always favourable. For some researchers and stakeholders, technological innovations in translation lead mostly to negative consequences, including the "automation" or "dehumanization" of the profession. Taravella (2013) notes that "a good number of the questions and concepts that make the literature on the establishment of information technologies and on the evolution

[23] *La contribution des non[-]humains ne peut plus être ignorée ou minimisée par les sciences sociales, car les investissements croissants dans la recherche et dans l'innovation technique en augmentent le nombre de manière exponentielle. [Cela dit,] [c]ette prise en compte du rôle actif des non[-]humains a de nombreuses conséquences.*

of the translation market are negative"[24]. Nevertheless, positive attitudes may also be found: "Collaboration and innovation [(Brunette & Gagnon, 2013)] or even productivity gains [(Zapata Rojas, 2012)(...)] are also at the heart of the discussion on the language industry" (*ibid*.).[25]

In any case, as Pym (2011, 5) puts it, "[t]he technology, for better or for worse, is here to stay. Few societies are able to refuse the use of a technology once acquired." Hence, it becomes crucial in TS research to consider both the negative and positive changes produced by new technologies in order to advance, as Bowker (2011) contends:

> By assessing and understanding these changes, we will be in a better position to develop strategies for moving forward, such as designing new types of resources for translators or making modifications to the curricula of translator training programs in order to reflect the new reality of the translation profession.

Most of the negative positions and attitudes towards technology could be thought to stem from the long-standing debate on a specific NLP application that has existed for almost as long as computers have existed: machine translation (MT).

MT and other applications will be discussed in the following subsection.

2.1.3. The Debate on Machine Translation, and the Emergence of Computer-Aided Translation Tools

Ever since the first experiments in NLP in the early years of computing, there has been a recurring debate on whether translation might become a fully automated task and,

[24] *Bon nombre des questions et des concepts qui parsèment la littérature en matière d'implantation de technologies de l'information comme en matière d'évolution du marché de la traduction sont négatifs.*
[25] *La collaboration et l'innovation [(Brunette et Gagnon, 2013)] ou encore l'amélioration de la productivité [(Zapata Rojas, 2012)(...)] sont également au cœur de la conversation sur l'industrie de la langue.*

subsequently, on whether the future of professional translators is threatened. This debate, however, has seemed to lead to the same conclusion: machines do not have the ability to perceive all subtleties of languages, whereas humans do (Arnold 2003); MT will never *totally* replace human translators.

In 1998, Shneiderman (1998, 293–294) stated that

> [s]tructured texts such as weather reports are translatable [by the machine]; technical papers are marginally translatable; novels or poems are not translatable. Language is subtle; there are special cases, contexts are complex and emotional relationships have a powerful and pervasive effect in human-human communication.

Thus, the realization that machines cannot process all subtleties of language has rather led – and will continue to lead – to significant efforts to develop technological applications that can aid human translators, that is, computer-aided translation (CAT) tools. Back in 1980, the visionary Martin Kay wrote (1980, 3–4):

> Translation is a fine and exciting art, but there is much about it that is mechanical and routine and, if this were given over to a machine, the productivity of the translator would not only be magnified but his work would become more rewarding, more exciting, more human [...] [I]f the need for translation is as great as it is said to be, the computer is our only hope.

Thus, rather than trying to replace human translators, computer tools ought to assist translators in efficiently and effectively carrying out their tasks (Bowker, 2002). But CAT tools have not all been created equal, and therefore lead to different challenges. Debates often tend to place these tools at the same level with MT and other LTs, as it will be discussed below.

2.1.4. The Debate on Computer-Aided Translation Tools, and the Investigation of Other Natural-Language Processing Applications and Emerging Interactive Technologies

As observed by the research community, CAT tools have significantly evolved over the decades and have modified the way human translators interact with machines to obtain and produce information (Bowker 2002). Online dictionaries and term banks, bilingual concordancers and translation memory systems (TMSs) are among the most widely used tools in our day. This being said, even though the primary goal of CAT tools is to assist translators by carrying out certain linguistic and computational tasks, a full integration of these applications into the profession has not yet been achieved, due to a number of technical and pedagogical challenges. Van der Meer (2011) argues that

> [t]echnology is often thought of as an answer to [the growing demand]. But along with the technology come many new challenges. It is simply impossible for a translator who is trained in the language arts to keep up with the technology. And if she tries, frustration grows when she finds out that translation tools do not really work together very well.

But challenges are difficult to foresee. Bowker (2011) notes that

> [w]hen new technologies are first introduced into a profession, it is not possible to know immediately what the long-term effects will be. It is only with time that the impact of applying a given technology can be observed.

Once technologies have been introduced and long used in different translation contexts, various conclusions and reactions emerge among researchers, practitioners and industry stakeholders. Anthony Pym (2011), for instance, questions the appropriateness of using certain technologies, arguing that these technologies may have an impact on the whole translation process and, consequently, on translated texts —and on language in general. He explains (2011, 2):

> Technology does not necessarily make things better or more efficient. Just as the internal combustion engine created traffic jams, if not global warming, so translation memories [TMs], along with [MT] engines and quick online documentation, can extend the list of alternatives only to impede efficiency in selection, undercutting intuition.

According to Pym, external aids or "memories" have a direct impact on the translation process and product because the translator's memory and other cognitive functions are altered, affecting thus the linearity and cohesion of the target texts. In short, technologies are believed to impact negatively the overall quality of the translation. This problem is attributed mostly to the extended use of TMSs (Macklovitch and Russell 2000; Bowker 2006; Garcia 2009; Pym 2011). Text segmentation, lack of cohesion and coherence, and inconsistent terminology are among the several drawbacks attributed to the use of TMSs in particular (Bowker 2006):

> To translate the overall message of [a] text, translators often need to work outside the artificial boundaries of sentences, so the sentence-by-sentence approach imposed by the TM may not be conducive to effective translation of the text's message as a whole. […] TMs may [also] hinder [terminology] consistency in some cases. A TM may contain texts that were written by different authors for different clients, and subsequently translated by different translators […] creating what is sometimes referred to as a terminology "train wreck".

Let us note also that the use of TMSs has been observed to lead to other transformations in the translation industry, particularly in the translator/employer and translator/client relationships. These transformations are also widely discussed in the TS literature. For instance, Garcia (2009, 201) points out that:

> TM became the interface between [language service providers] and freelance translators, allowing them to collaborate in large-scale translation projects. TM […] came into its own with localization. Ownership of and proficiency with an industry-compatible TM software suite soon became indispensable for aspirants to this kind of work. […] Over time, however, what had commenced as a translator's tool became something that language vendors imposed on their pool of freelancers, and finally—once major translation buyers became aware of the benefits—was in turn imposed on language vendors by corporations.

More recently, however, as Garcia (2009) notes, although TMSs are still widely used, the translation sector seems somehow to be moving beyond them. A reason for this may be TMSs' limited capacity to address the bulk of translators' genuine needs at the digital age (*ibid*.). Although the advantages of TMS are important, the many disadvantages from the users' perspective remain significant (LeBlanc, 2013).

Now, if MT and TMSs have dominated the majority of the debates on technologies and translation, other applications have had a rather positive impact on translators and on the profession, and have pointed towards promising results. Thus, newer applications and emerging interactive modes have been slowly changing the negative attitudes towards technologies among the TS community, as I will discuss now.

Recently, investigations in CAT have paid greater attention to the end users of technologies, and have also taken into account the potential of emerging input modes such as voice and stylus, and innovative approaches such as collaborative translation and crowd-sourcing.

An important element in these innovations has been MT, but this time with human intervention. This latter approach is commonly known as post-editing (Barrachina et al., 2009; Foster, Isabelle, & Plamondon, 1998; Koehn & Germann, 2014; Koehn & Haddow, 2009). In post-editing, the goal is no longer to produce translations in a totally automatic manner, but to assist the translator (or the post-editor) in building a human translation with the "least effort possible". This approach is *arguably* becoming a well-accepted practice in the translation sector since, according to some experiments' results, it allows for larger volumes of translations to be produced and for time and costs savings (Bowker & Buitrago Ciro, 2015; Mesa-Lao, 2014, p. 99). In a traditional post-editing workflow, the user replaces, deletes, inserts or moves pieces of text (i.e., the MT output) in a given text editor or text field using the keyboard and occasionally the mouse (Alabau & Casacuberta, 2012).

It is worth mentioning another major development in NLP that seeks to provide translators with assistance to improve performance and to reduce the technical effort:[26] interactive translation prediction (ITP). ITP systems provide target-language equivalent suggestions coming not only from MT (such as in Langlais et al. (2000)'s and Macklovitch (2006)'s TransType and TransType2 systems), but also from any other type of bilingual resource (e.g. a TM, a bilingual dictionary, a catalogue of bilingual sentences [in a two-column format in a MS Excel spreadsheet, for instance]) (Pérez-Ortiz, Torregrosa, & Forcada, 2014). In ITP, a possible solution pops up in the text field once the user has typed enough information for the system to come up with

[26] Technical effort refers here specifically to the physical act of producing a text in a KM environment; the number of keystrokes (and mouse events) needed to type a text. See discussion in chapter 6.

a suggestion. When using such systems, significant savings in the number of keystrokes and mouse activities and, potentially, time, can be theoretically achieved (*ibid.*). ITP features have been added to post-editing environments and some pilot experiments have reported on these combined approaches (see Carl et al. (2016)).

Now, emerging (non-KM) input modes, particularly speech input, are also being explored in TCI (Alabau, Rodríguez-Ruiz, Sanchis, Casacuberta, & Martínez-Gómez, 2011; Barrachina et al., 2009; Garcia-Martinez et al., 2014; Mesa-Lao, 2014). The results of these efforts have been rather positive. In the context of post-editing environments, for instance, even though keyboards seem to be the easiest input method when only minor editing is needed, as García-Martínez et al. (2014) put it,

> when the text requires major changes (e.g. editing larger segments of text), typing could be optimized using other input modalities. Moreover, if the post editor is not a touch typist, then she has to switch visual attention back and forth between the screen and the keyboard making the task more complex.

Thus, speech input offers a third dimension to the post-editing task, making it possible to combine different input modes or to alternate between them according to the difficulty of the post-editing task and to the changing conditions of human-computer interaction.

Furthermore, stylus input has also been explored for post-editing purposes (Alabau et al. 2010; Alabau, Sanchis, and Casacuberta 2014). In this approach, post-editors would use, for instance, standard proofreading handwritten gestures for substitutions, deletions, insertions and transpositions (see Figure 6 on page 85). This is what Alabau et al. (2011) call "online" handwritten text recognition (HTR), that is, the recognition of a user's handwriting input directly on a touch-enabled device.

HTR is a promising recognition-based technology to be explored in TCI. Research on HTR first emerged as an effort to transcribe text images (e.g. digitalized copies of manuscripts), a task that can become computationally difficult when the writing style is too variable (continuous, cursive handwriting, presenting skewed/slanted words; irregular calligraphy; etc.) or, in the case of old manuscripts, the original copy is particularly degraded. With online HTR ("online", because the recognition is done as the user handwrites on the interface's touch-screen with a digital pen (or stylus)), it is now possible to investigate various ways to integrate a translator's handwritten gestures not only to proofread MT output but also to carry out other translation-related tasks, particularly during preparation of a translation (see below), and even revision of a translated text (Alabau & Leiva, 2014) (see Chapter 4).

Recent results on the cognitive and technical advantages of VTS interfaces seriously challenge the primacy of KM GUIs, and encourages simulation and prototyping of new multimodal environments that incorporate voice-and-pen input in our education and professional lives (Oviatt, 2012, p. 423).

With this in mind, I carried out a formal pilot experiment within the framework of a doctoral seminar[27] to explore handwritten interaction in source-text preparation for TD.

Source-text preparation tasks have been under-investigated in TS, including how translators annotate source text copies prior to producing the target texts. There is no “right” or “wrong” way to prepare a source text; it is rather an intuitive action, which depends on the translator’s level of experience and/or level of comfort with the source and target languages and knowledge of the terminology and subject field, and on tool availability. Traditionally, a translator would have the source text on paper to make annotations (e.g. to write down equivalents) or drawings such as circles, arrows or symbols in preparation for the written or oral production of the target text. The preparation step is in fact very important in the translation process. According to Gouanvic (1976, p. 252),

> regardless of the device used [to type or to dictate a translation], it is essential to carry out a very detailed research of unknown terms, expressions, etc. or the particularities of a text (including difficult passages), before the final writing or dictation phase.[28]

[27]The experiment was carried out in Winter 2013 within the framework of the *TRA6985 Current Developments in Translation Studies II* seminar. The experiment was performed under the supervision of Professors E. Marshman and J. Quirion. The results of the experiment were presented at a graduate student conference at the University of Ottawa in April 2013, but were not published in the form of a paper. The methodology and the results of the pilot experiment are presented in Appendix C of this thesis. The research ethics board notice of approval is provided in Appendix D.

[28] *Quels que soient les appareils utilisés [pour saisir ou pour dicter une traduction], il est essentiel d’effectuer une recherche très soigneuse des termes, expressions, etc. inconnus, ou des particularités du texte (y compris des passages obscurs), avant la phase finale de la rédaction ou de la dictée.*

Thus, such crucial task may be simulated using a stylus-enabled touch-screen device. In addition, the stylus can also be used as "direct" pointing device on the interface. This characteristic of the stylus has made it attractive for software developers for many years, since it allows more direct functionalities than passive, indirect pointing devices (Lee, Isenberg, Riche, & Carpendale, 2012; Shneiderman, 1998, p. 318).[29] Source-text preparation is a whole topic in its own, and is worthy of a larger-scale study, which is beyond the scope of this work. The results of this pilot experiment also add to the design of an ITD environment's UI, aligning thus with the objectives of this thesis, but a larger-scale experiment will need to be carried out in future work.

Lastly, it is important to note that one of the main motivations for exploring stylus interaction is mobility (Alabau & Casacuberta, 2012; O'Brien, Moorkens, & Vreeke, 2014). Using a mobile, touch-enabled stylus interface would it possible for translators to carry out certain tasks more easily than keyboard-based interfaces while commuting or travelling for instance (Alabau & Casacuberta, 2012). Apart from the advantages of mobile use, stylus interfaces are considered a promising avenue for minimizing cognitive load and supporting human cognition in general (Oviatt, 2013).

As has been discussed so far in this chapter, innovation in translation research and CAT-tool development has continued over the years, but attitudes towards technological developments have been divergent. As I will argue and demonstrate in this thesis, for new developments to succeed, greater consideration will need to be

[29] The mouse is an example of an indirect, passive input mode. For instance, controlling a pointer with an external device and moving one's fingers to click or scroll down are not concrete indicators of the user's actual intentions. Unlike the mouse, stylus and finger input are considered direct, active input modes.

given to the translation process as a whole and to the end-users of technological applications.

2.1.5. Changing Attitudes towards Technology: Applications and Devices for Improved Human Cognition and Well-Being

Researchers have argued that the CAT tools marketed and used today have rarely been designed with the end user in mind (LeBlanc 2013; Taravella and Villeneuve 2013; O'Brien 2012; Carl, Dragsted, and Jakobsen 2011); taking into account the human factor will be essential in future CAT-tool research. O'Brien (2012, 109–111) questions:

> Can translators make their jobs more human through HCI? Can we allow the machine to take over the boring, repetitive tasks and free ourselves up for the harder, subtler and more complex problems? And what are those problems that machines cannot solve, but human translators can? These are some of the large questions facing us.

Among emerging technologies, speech-and touch-based applications represent a promising avenue to achieve a better interaction between human translators and computers. Jurafsky and Martin (2009, 8) argue that humans have a certain predisposition to accept machines as social entities, which explains the great interest in speech technologies, for they allow interaction through natural language:

> It is now clear that regardless of what people believe or know about the inner workings of computers, they talk about them and interact with them as social entities. People act toward computers as if they were people; they are polite to them, treat them as team members, and expect, among other things, that computers should be able to understand their needs and be capable of interacting with them naturally […] Given these predispositions, speech- and language-

> based systems may provide many users with the most natural interface for many applications.

Today, with the introduction of smartphones, tablets and other MIs in which speech is one of the main input modes, there is an excellent opportunity to reinvent the way translators interact with different tools to prepare, produce and revise translations, and even the way they interact with their peers, employers and clients. MIs modify not only the way humans interact with machines, but also the way they interact with the information they need, and with other people. Lastly, MIs support human cognition, ideational fluency and well-being (Oviatt, 2012) (further discussion is provided in Chapter 4).

Now, having discussed TCI from different angles and lent support to the idea of exploring VR and multimodal interaction in translation, let us look at some notions in HII and discuss how this field can inform and strengthen TII research.

2.2. Translator-Information Interaction in the Digital Age

The study of the interaction between humans and information is not new. However, with the advent of ICTs and, in particular, of the Internet, the field of HII became popular within the research community in information science, computer science and other disciplines (Fidel, 2012, pp. 46–48; Marchionini, 2008), including TS. Some researchers even argue that HII goes well beyond HCI, and that the latter will soon become an “old-school” term, since “the interaction is with the information. The information we need and use is becoming pervasive and ubiquitous, and the computer just happens to be one of the mediums that supports or mediates the interaction with the information” (A. Parush, personal communication, March 7, 2013).[30] Indeed, humans are in constant interaction with information, be it through machines or not. Any contact with the universe through one or several senses is an interaction with information, from the moment we wake up until the moment we fall asleep, and even in our dreams (Fidel, 2012, p. 1). These realization has also led to a number of developments on haptic technologies (Stone, 2001) as well as embodied and tangible interaction (Dourish, 2001; Hornecker & Buur, 2006).

Information is everywhere (Kurzweil, 2013, pp. 2–3). This being said, today, the massive influx of digital wireless/mobile devices has led to new ways of accessing enormous quantities of information and services at any time and from practically anywhere. Aubert, Cohendet, and Montreuil (2010) describe how ICTs have dramatically evolved during the past 40 or 50 years. These technologies, as they explain (2010, pp. 8–9), have led to

> profound changes affecting individuals and organizations as much as the society in general […]. [They] have made information access constant, transparent and more and more complete. Indeed, the challenge is no longer to access information, but rather to filter pertinent information.[31]

But the 1990s brought radical changes (Barabasi, 2012, p. 38),

> a massive influx of data coming mostly from the Internet, from telecommunications and from digital technologies in general. All

---

[30] Avi Parush is a professor (on leave as of June 2016) at the Department of Psychology at Carleton University and teaches at the Master's program in HCI.

[31] *Des changements profonds touchant autant les individus et les organisations que la société en général […]. [Elles] ont rendu l'accès à l'information constant, transparent et de plus en plus complet. En fait, le défi n'est plus d'accéder à l'information, mais bien de filtrer l'information pertinente.*

> disciplines […] have had to face this explosion of information to be managed […]. Today we are living in the age of data.[32]

ICTs are here to stay and have become more and more accessible to anyone, anywhere. It even becomes difficult to define them and to differentiate them: for example, mobile phones (or smartphones) are not only telephones; they contain also, among many other things, personal assistants, web browsers and mobile TVs (Harper et al., 2008, pp. 34–38). These interactive devices offer all the advantages of mobility and multimodality (Schalkwyk et al., 2010); this is why the interest in mobile and cloud computing has increased in recent years, since they represent the future of technology—and therefore of translation technology, as I argue in this thesis.

A current subject in the era of ICTs, the study of the interaction between humans and information is, by nature, complex and multidisciplinary (Fidel, 2012, pp. 17–21). According to Marchionini (2008, p. 171), HII is complex because it implies considering every element and every aspect of the informational work: the interaction process and the changes that result from that interaction at the level of the individuals searching for information and the objects or systems used to obtain the information.

Two sub-fields of information science are particularly well-grounded today, and seem to offer great potential in TII: information retrieval (IR) and information behaviour (IB). The former investigates the models and mechanisms of computer systems that

[32] *un afflux massif de données provenant d'abord d'Internet, des télécommunications et des technologies numériques en général. Toutes les disciplines […] ont été confrontées à cette explosion d'informations à gérer […]. Nous vivons désormais à l'ère des données.*

allow or facilitate information research. The latter examines information seeking strategies, information evaluation criteria, and the modalities and contexts of information use. In other words, while IR focuses on developing and improving informational tools, IB investigates the ways of browsing the different sources of information, and of evaluating if the information found is adequate for solving a given problem, in order to use it according to the constraints (Fidel, 2012, pp. 35–37) imposed by the context. Fidel (2012, p. 85) explains that context is important because, even before carrying out any search, it shapes the informational needs, since the motivation to search for information is not only cognitive, but also contextual.[33]

In sum, HII and in particular IR and IB offer a great potential in TS research since, in the search for translation tools that are more efficient, it is necessary to meticulously investigate translators' behaviours with the different informational resources they use to carry out translation tasks. Hence, more development on those fields will be needed in future work in TII.

2.2.1. The Internet as a Primary Source for Information Retrieval

The Internet has become the primary source for information retrieval in translation (Borja, 2008; Simard, 2013). Very few translators still take the time to open, or even to carry along their (huge) paper dictionaries, term records and language books, to name only a few traditional informational resources. On the web, translators can find hundreds of monolingual and bilingual dictionaries, concordancers and

[33] Contrary to other studies in HII, taking context into account is the basis of Fidel's reflections (2012, pp. ix–x), which explains the subtitle of her book "*An Ecological Approach to Information Behavior*".

biconcordancers, terminology databases, grammar and conjugation guides, encyclopaedia and other documentation; in sum, practically all the information that may be useful when producing a translation. Borja (2008, p. 252) explains:

> In just a few years, the Internet has become the main resource for terminological and conceptual documentation for professional translators, and today it can be said that 95% of the material they need can be found on the web.

Indeed, the Internet is considered the informational resource *par excellence*, the "El Dorado" of knowledge (Duval, 2012, p. 50). Now, the fact that it is becoming accessible from practically anywhere and anytime (see definition and examples of cloud computing in footnote 2 on page 3) leads humans to develop new behaviours and new ways of interacting with information; of understanding, using and producing information.

As it is the case in many contexts and domains, the Internet is bound to bring radical changes to the translation field, as Garcia (2009, p. 211) explains:

> As Internet becomes a true utility, translation is not the only profession to experience the stress of the digital age. Translators will still be needed, but their working conditions into the next decade will be quite dissimilar to those of [previous decades].

Thus, it will become necessary in TS to examine the impact of ICTs, particularly the Internet and Internet-based information, on the translator's cognitive processes. In a cognitive psychology study, Duval (2012) explains that some cognitive functions such as reading, learning and memorizing are affected by the (intensive) use of the Internet. In fact, people will turn to a search engine to search answers to even the simplest question. Just knowing that a piece of information is readily available anywhere and anytime leads humans not to memorize it. Furthermore, with the introduction of tablets and smartphones, human beings had never been so exposed to texts and information; the reading task is thus believed to have evolved. Reading is no longer just decoding words and understanding a message, but also, for instance, deciding if one should click on a given hyperlink, or not.

As the Internet becomes more and more available to humans, it will be an inexorable object of research in TS. The UNESCO Broadband Commission estimates that by 2017 half of the world's population will have access to the Internet.[34] Indeed, great efforts are made by many individuals and organizations to continually increase and facilitate access to the web in order to fill the "digital gap", that is, the distance between those who have full access to ICTs (the "information rich") and those who have limited or no access (the "information poor") (Havens & Lotz, 2012, p. 27). For example, for a number of years, several city governments, media organizations and citizen groups have been lobbying for free and unlimited access to a broadband Internet network for every member of their communities. Today, in most of the world's major cities, wireless Internet networks are available in trains, buses, public transportation stations, airports, coffee shops, restaurants and other venues. Furthermore, wireless Internet has been recently made available on airplanes while in flight at over 10,000 feet within certain coverage areas.[35]

---

[34] Source: http://www.broadbandcommission.org/Documents/reports/bb-annualreport2014.pdf. Retrieved 02-05-2016.

[35] This sentence was dictated in late 2015 from an airplane during a domestic flight within the United States, using a cloud-based SI VR system accessed from a tablet.

Perhaps the most ambitious initiatives for worldwide Internet connectivity are those commenced in 2013 by two ICTs giants: Google's *Project Loon,*[36] and Facebook's *Internet.org,*[37] which aim at providing the world's entire population with wireless Internet access. Facebook's founder Mark Zuckerberg (2014) explains:

> Connecting the world is one of the fundamental challenges of our time. When people have access to the Internet, they can not only connect with their friends, family and communities, but they can also gain access to the tools and information to help find jobs, start businesses, access healthcare, education and financial services, and have a greater say in their societies. They get to participate in the knowledge economy.

To achieve its goals, in addition to building drones and satellites which will beam Internet signals to the Earth, *Internet.org* plans to lower the costs of connectivity, to build infrastructure adapted to the particular needs of different communities, to improve the efficiency of data networks and services and to develop business models that give people new ways to go online, to obtain information and to connect with other people.

Google's approach is just as ground-breaking and ambitious: their goal is to provide Internet access to the entire planet by putting out a ring of balloons that fly around the planet, at the stratosphere level, beaming Internet connectivity to a ground area of 40 km in diameter.

[36] Source: http://www.google.com/loon/. Retrieved on 23-11-2015.
[37] Source: http://internet.org/. Retrieved on 02-05-2016.

These examples show a great tendency for the Internet to become the world's primary information resource or, as Zuckerberg puts it, the "backbone of the knowledge economy"[38].

Thus, without a doubt, the Internet will play an increasingly important role in TII research, both for understanding translator-information behaviours and for improving the quality of the information and of information research tools used by translation professionals. For these reasons, within the framework of the present study, only Internet-based informational resources will be investigated since the web is and will be the main source of information for translators in the digital age (Borja, 2008; Garcia, 2009; Simard, 2013).**Concluding Remarks on Translator-Information Interaction**

Several questions may motivate TII research. For instance, how well do human translators work with the information and informational tools they currently have at their disposal? How accurate, rich and relevant is the information they find? How user-friendly are informational tools for translators? How can the information and the tools be improved to maximize translators' efficiency and well-being, and the quality of their work? More specifically, for example, what are the advantages and drawbacks of integrating informational tools to a translation environment such as a TMS or an ITD environment, as compared to having an array of external web-based resources? In an effort to find answers to such questions, a number of scholars in the TS field, and also in Information Science, have undertaken studies of different kinds involving different tools and different data-collection methods. Examples of such

[38] Zuckerberg narrates the video "Making the Internet Affordable": http://www.youtube.com/watch?v=AdXwthh-xLQ. Retrieved on 23-11-2015.

investigations include Sales (2008), Désilet et al. (2009), Sales and Pinto (2011), Simard (2013), Palomares Perraut and Gómez Camarero (2014), Ehrensberger-Dow and Massey (2014), Enríquez Raído (2014), and Volanen (2015).

The field of TII is still at its emerging state, but will undoubtedly become more and more important in TS. In fact, in the search for tools that are better adapted to the translation workflow and to the informational needs of translators, it will be essential to investigate the way translators interact with information and the information tools currently available to them.

The following chapter provides a close-up look at current translation techniques, while examining the advantages and drawbacks of these techniques and supporting the need for change and innovation in TCI and TII.

Chapter 3 — A Close-Up Look at Current Translation Techniques: From Typed Translation to Translation Dictation with Voice Recognition

"To me, the greatest pleasure of writing is not what it is about, but the music the words make." — Truman Capote

This chapter provides a discussion on current translation techniques, from typed translation (TT) to TD with VR (TDVR). First, it provides a general background on the use of written and oral techniques, dictation tools, and VR technology in translation. Then, it examines the advantages and drawbacks of these techniques and supports the need for change and innovation in future developments in TCI and TII.

3.1. Background

Translation has existed for millennia. Archeological evidence (tools, cave paintings, carvings, etc.) suggests that early humans, who made their appearance hundreds of thousands of years ago, had the type of intellect that could support language. This ability to produce speech and to communicate is in fact what distinguishes us, *Homo sapiens*, from other species (O'Grady, 2004, p. 1). Now, (oral) translation activity is believed to date back to the development of humans' ability to produce speech and use language. Consequently, written translation very likely emerged in parallel with alphabets, writing systems and writing tools. The Rosetta stone, displayed at the British Museum in London, is one of the oldest standing artifacts that account for translation activity in ancient times (Budge, 1989). Throughout history, translators, as well as writers of any kind, have adopted different tools according to ongoing developments in science and technology: from stone-engraving tools, to papyrus and ink pens, to typewriters and – most recently – PCs.

Yet the physical act of writing is not very popular among certain translators and writers in general: many wish that their hands moved at the speed of their thoughts. In fact, either by preference or because of health issues, some writers and translators have opted for dictation, i.e. the oral production of their texts into a voice recorder – also known as a dictaphone – for later transcription, or even directly to a secretary or professional transcriptionist, who transcribes the words as they are being dictated (Daiute, 1985; Gingold, 1978; Gouadec, 2007, pp. 363–364; Gouanvic, 1976; Hétu, 2012; Jiménez Ivars & Hurtado Albir, 2003; Jiménez Ivars, 1998, pp. 179–181; Laroque-Divirgilio, 1980, 1981). Besides, it is known that dictation was the only means for some famous writers to convey their thoughts: the blind John Milton dictated the epic poem "*Paradise Lost*" to his daughters, and Henry James dictated his later novels after a repetitive stress injury (Jurafsky & Martin, 2009, p. 285). In sum, writers who can speak their ideas (to transcriptionists or into a recording device) have historically had the quickest means of getting words on paper (Daiute, 1985, p. 31).[39]

TD[40] was a very popular –and indeed "efficient and effective" (Gingold, 1978)– working technique in the 1960s and 1970s, but started to fade away as professional

[39] Note that the time and human resources needed to have orally-produced texts transcribed needs to be taken into account when assessing the efficiency of dictation, and when considering adopting this technique, the way it was used decades ago. See below for a more detailed discussion.
[40] See the following section for more details on the various forms of TD, including TD with VR.

translators' workstations experienced the massive influx of typewriters and PCs (Laroque-Divirgilio, 1980, 1981). Nowadays, TD is rather an uncommon practice because translators have the ability to type their translations directly into the computer (a technique that will be called typed translation [TT] in this research), thus "saving" the extra time (and money) needed to transcribe their dictations (or to have them transcribed by a third party). This being said, with the increasing global demand for top-quality translation, which currently exceeds translators' (and MT engines') capacity (Barabé, 2013), things are bound to change—yet again.

The tremendous improvements in natural language processing (NLP), and particularly in VR technology, over the past two decades, provide an excellent opportunity today to bring back TD to the profession; this time in a new form, using emerging technologies and applications. Some TS scholars even believe that there should be a full (re)integration of TD to translator training programs and an optimal adaptation of VR technology to translators' tasks (Ciobanu, 2014; Gouadec, 2007; Mees et al., 2013). Gouadec (2007, p. 363), for instance, argues that

> [i]n all probability, the translation industry in the twenty-first century will […] go oral. [VR] and dictation software is probably going to radically change the way translation is perceived and practiced […]. Dictated translation[41] will become the norm once again.

In fact, despite their error rates and limitations, off-the-shelf SA VR systems have been part of certain translators' toolbox for over a dozen years now (Bowker, 2002, p. 42; Ciobanu, 2014; Gouadec, 2002, p. 123; Zapata Rojas, 2012, chapter 3). In many cases,

[41] Dictated translation and TD refer to the same notion.

translators who use VR software are those who once dictated with the aid of dictaphones and/or transcriptionists back in the 1960s and 1970s (Seaman, 2002).

Historically, translation has been an important part of the development in VR technology. The first efforts involving both VR and translation in the early 1990s aimed at providing conversational agents (computer systems with which the user can have a conversation) (Weizenbaum, 1966) with the ability to translate, thus allowing for simple conversation between people who do not speak the same language. Such "artificial interpreters" (or "voice-to-voice translator" applications, as commercialized today with mobile devices) use VR to process the user's input in a language A, a machine translation (MT) system to automatically produce a translation into a language B, and speech synthesis to emulate human speech in the language B. The interlocutor can then speak their response into the system in language B, and the system translates and "responds" in language A, and so forth.

In the mid-1990s, efforts to adapt VR technology to human translation took place for the first time. Such developments focused on minimizing word error rates by combining VR and MT (e.g. the TransTalk project described in Dymetman et al. (1994) and Brousseau et al. (1995)). Hybrid VR/MT systems have access to the source text and use probabilistic MT models to improve VR. Dymetman et al. (1994, p. 2) explain:

> We do not try to replace the human translator by a machine (a hopeless endeavor, in general), but undertake instead the more realistic task of providing a dictation tool to the translator. Our aim is to use [MT] to make probabilistic predictions of the possible target language verbalizations freely produced by the translator and to use these predictions to reduce the difficulty of the [VR] task to such an extent

> that complete recognition of the translator's utterances can be achieved.

In practice, in an English-to-French translation context, if the hybrid VR/MT system must choose between two acoustically similar French words such as "*chevaux*" and "*cheveux*", the presence of the word "horses" in the English source text—which is pre-processed by the system—would guide the system in making a decision and producing an accurate transcription (Brousseau et al., 1995, p. 193). Similar efforts in this vein have continued over the years, highlighting the various challenges of VR/MT integration (Vidal et al., 2006; Désilets et al., 2008; Reddy and Rose, 2010; Rodriguez et al., 2012), and the potential benefits of using speech input for human translation purposes (García-Martínez et al., 2014; Mesa-Lao, 2014). Likewise, further efforts have been made by translation trainers and researchers to evaluate the performance of students and professionals when using off-the-shelf VR systems for straight TD (Dragsted et al., 2009; Dragsted et al., 2011; Mees et al., 2013) and to introduce them to this technology (Romero-Fresco, 2012); and to assess and analyze professional translators' needs and opinions vis-à-vis VR technology (Ciobanu, 2014; Zapata Rojas, 2012).

In sum, research in VR for translation purposes has been around for many years, but has been rather hesitant due to a number of challenges of different nature. Challenges have included the inability of VR systems to flawlessly transcribe continuous speech back in the 1990s (De Schaetzen, 1995, p. 685; Kurzweil, 2013, pp. 72–73, 122–123), the inability of today's commercial VR systems to process two or more languages at a time[42] and to carry out translation-specific tasks (Gouadec, 2002, p. 133), as well as the lack of TD- and VR-related courses in translator-training programs (Benis, 2002; Gouadec, 2007, p. 286; Mees et al., 2013), also essential for succeeding with the technology. On this latter issue, Mees et al. (2013) explain:

> The readiness to embrace new technologies is an important prerequisite for succeeding with the SR program. […] Universities have an important role to play in introducing SR in addition to other technologies as part of the translator's toolbox. For too long universities […] have engaged in traditional approaches to translator training, and have not addressed the changing preferences of students for obtaining and communicating information. [...] With the recent development of voice activation in [sic] smart phones and computers there is a golden opportunity to capture students' interest. As speech becomes a more reliable and popular input method, using SR may grow into a more natural working mode also in translation.

All things considered, studies on the integration and the adaptation of VR technology to the translation profession have been rather optimistic and researchers and developers point out the need for more efforts in this vein.

To sum up the current state of affairs, VR technology offers considerable potential for translation students and professionals, but it will need to be used differently and integrated more efficiently. One possible way to do this is by enabling multimodal interaction. In a recent study by Mesa-Lao (2014), student translators surveyed reported that they would welcome the integration of voice input as one of the possible input modes for performing translation tasks (in the case of Mesa-Lao's study, for post-editing MT output).[43] Indeed, 12 out of 15 respondents stated that they would

[42] Bilingual VR will be discussed more in depth in chapter 4.
[43] More details on MT and post-editing will be provided in Chapter 2.

definitely consider adopting VR in combination with non-speech input modes in their daily work as translators. VR is thus a complement rather than a substitute for non-speech input methods. Mesa-Lao concludes that further studies still need to be conducted in order to build up new knowledge about the many ways in which VR technology can be adapted to translation tasks. My research on ITD is thus a promising avenue, since it promotes the use of VR technology in a multimodal environment.

Let us now take a close-up look at current translation techniques, from TT to TDVR.

3.1. A Taxonomy of Translation Techniques

Today, there are two main techniques used in translation teaching and practice: typed translation (TT) and sight translation (SiT), as shown in Figure 1 below:

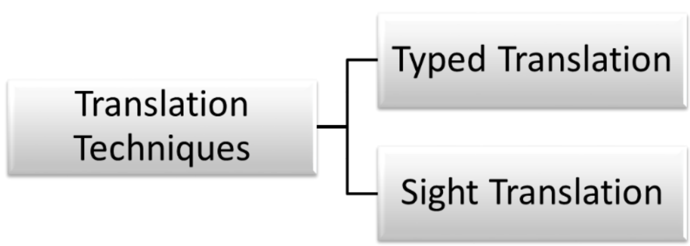


Figure 1. Translation techniques.

TT – the input of translated texts using a computer keyboard – is probably the most widely used technique today. Ever since the massive influx of PCs, translators have mainly typed their translations using the keyboard and a word-processing application. But TT is not the only means to get a translation “on paper” (or screen). As discussed in the previous section, many translators have used oral translation techniques such as translation dictation (TD), mostly in the 1960s and 1970s; in most of the cases, those oral techniques have proven to be more efficient, effective and ergonomic than TT (Gingold, 1978; Hétu, 2012; Laroque-Divirgilio, 1981).

TD is a form of sight translation (SiT). SiT – the oral rendering in a target language of a written text in a source language – has been vastly used for several decades mostly in interpreter training, but also in translator training and translation practice (Dragsted & Hansen, 2009; Dragsted et al., 2011; Gouanvic, 1976; Jiménez Ivars & Hurtado Albir, 2003; Jiménez Ivars, 1998; Lambert, 1989, 1991, 2004; Mees et al., 2013; Pratt, 1990; Spilka, 1966). SiT is divided into four types: (1) first-sight translation, (2) SiT with preparation, (3) consecutive SiT, and (4) TD. The product of the former three types is an oral translation (i.e., the end user of the translation only “listens” to the translation); the product of the latter is a written translation, as illustrated in Figure 2:

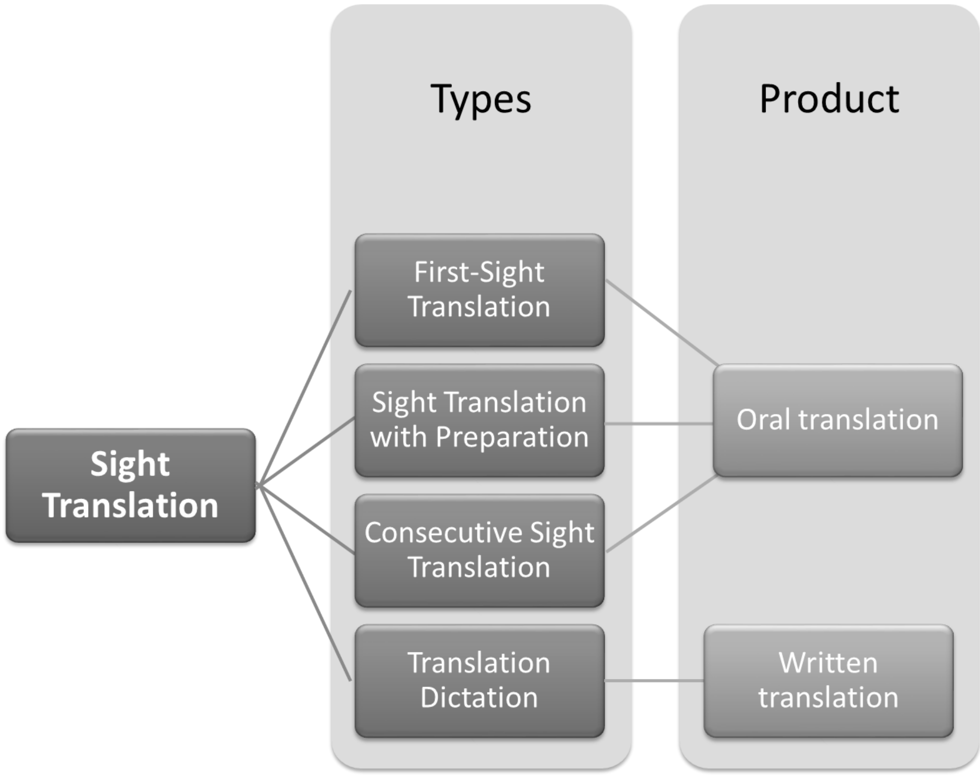


Figure 2. Sight translation. Types and nature of the product in each type.

Figure 3 below shows the different types of TD and the recipient of the dictation in each type. First, there is conventional TD, which can be performed directly in front of a transcriptionist or, most commonly, is recorded into a dictaphone for later transcription. Then, there is TDVR, which is performed using a VR system embedded in a PC.[44] (Lastly, there would be ITD, in which the transcription of the dictation is carried out primarily by a VR system in a multimodal environment augmented with

[44] One of the main characteristics of VR systems for PC is that they are mostly SA. This means that the user has to create a user profile and train the software to his or her voice. Then the system adapts dynamically to the particular voice and speech characteristics of the user. See Woodland (2001) for a review of adaptation techniques. In addition, as Leijten and Van Waes (2005, p. 741) argue, the performance of the SA system improves as (proper) usage and training time increase; in other words, the system "learns" from the user.

other input modes in addition to voice. Further discussion on multimodal interaction and ITD will be provided in Chapter 4).

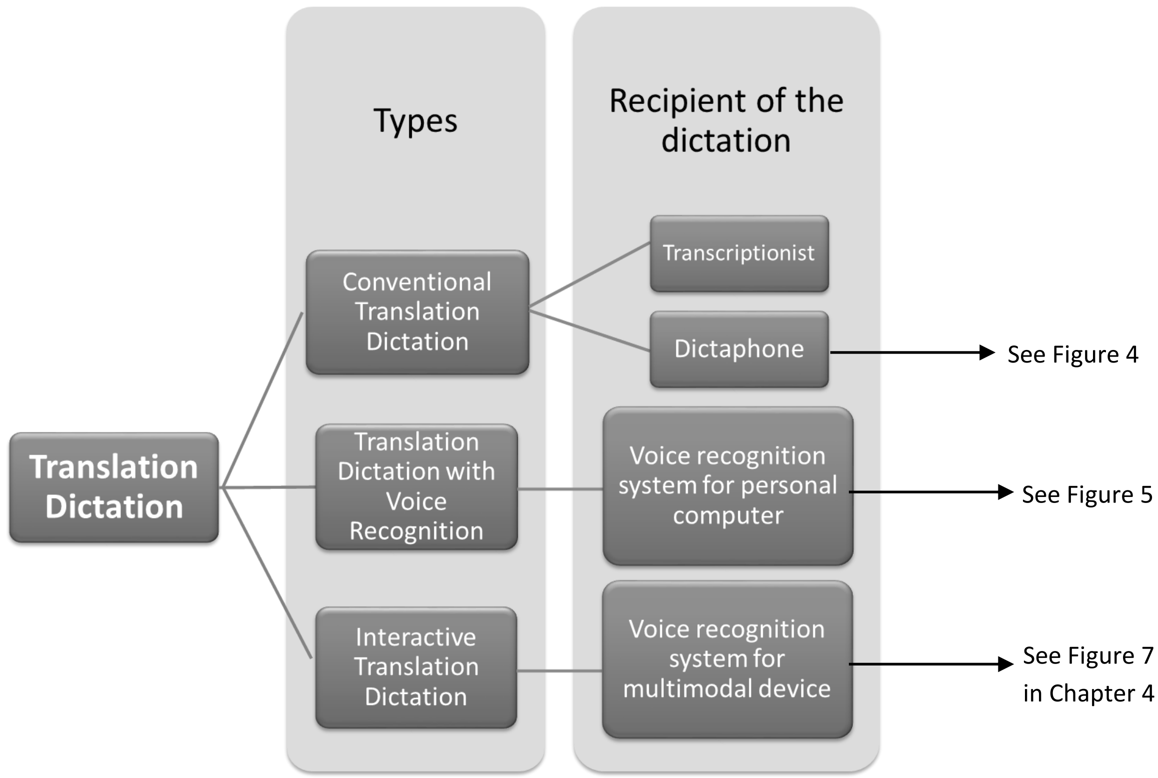


Figure 3. Translation dictation. Types and recipient of the dictation in each type.

Figure 4 below illustrates conventional TD with dictaphone (CTDD). Let us note that conventional TD directly to a transcriptionist is excluded from this classification because, even though used in some instances, it is rather an uncommon practice today and not ideal for achieving a good performance and cost-effectiveness (Y. Martin, personal communication, September 4, 2013).[45] In CTDD, the transcription of dictations is carried out by the client, by the translator him or herself or, most

[45] Y. Martin is a retired Canadian translator who worked using CTDD for over 40 years. I performed an informal experiment, including an interview, with this person at the early stages of my doctoral project.

commonly, by a professional transcriptionist, who listens to the recording and types it into a word processing application. In some cases, the transcriptionist may also use an off-the-shelf VR system to transcribe the translation by "respeaking" the recording (i.e., by listening to the dictation and simultaneously repeating (or "shadowing") it into the VR system).

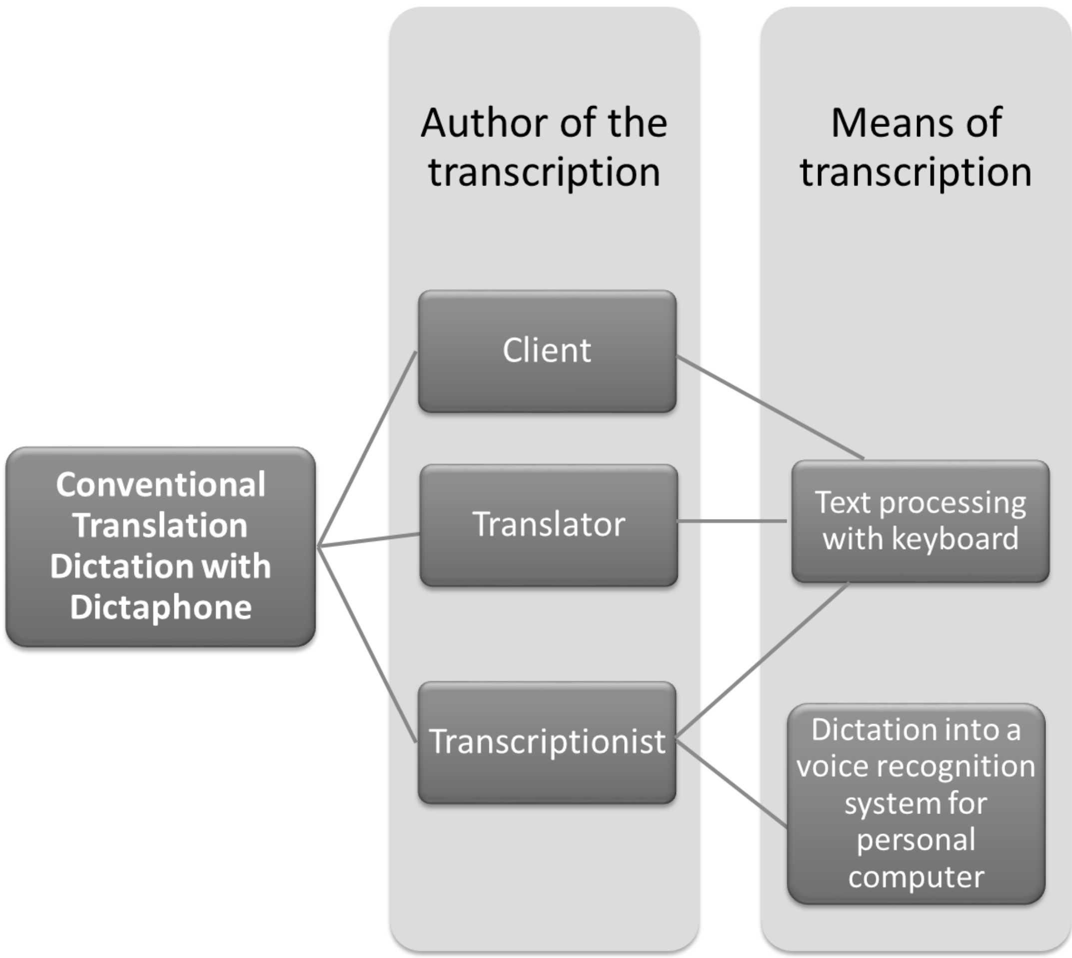


Figure 4. Conventional translation dictation with dictaphone. Authors of the transcription and means of transcription.

Figure 5 below shows the two types of TDVR, which are characterized by the different means of obtaining transcriptions of dictated texts. In real-time TDVR, the dictation is performed directly into a VR system, which transcribes the dictation simultaneously. In other words, the transcription appears on the screen as the user speaks, with a very short delay (usually within seconds). In offline TDVR, the dictation is recorded first using a dictaphone, and then the user transfers the audio file to the computer and sets the VR system to transcribe it in a text-processing application.

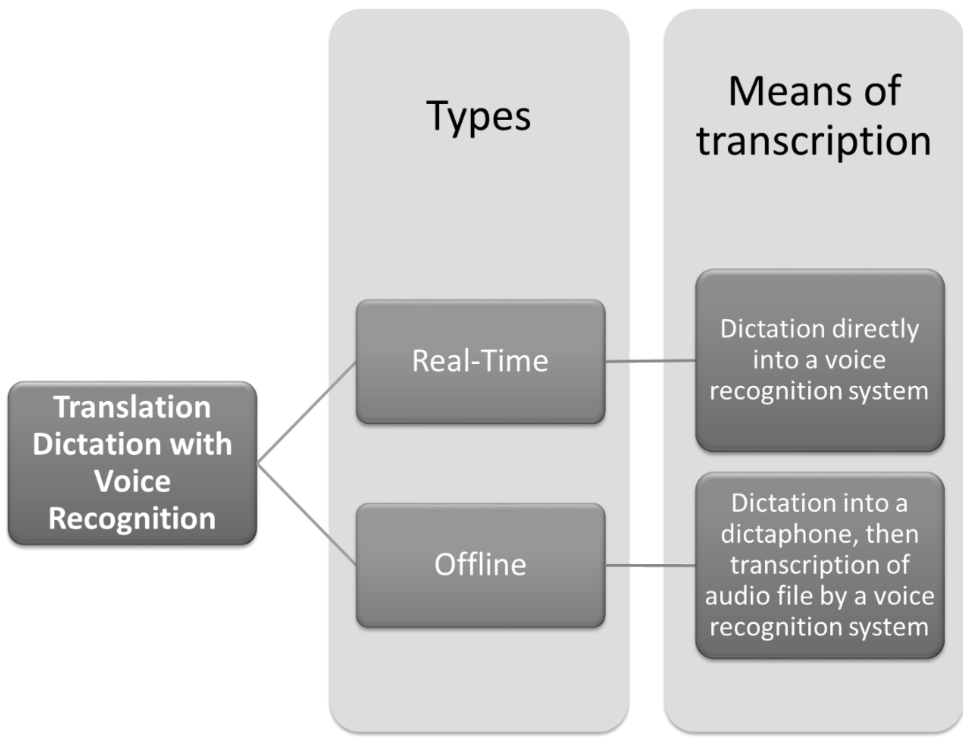


Figure 5. Translation dictation with voice recognition. Types and means of transcription.

The various advantages and challenges of current translation techniques are discussed in the following sub-section.

3.1.1. Advantages and Challenges of Current Translation Techniques

Current oral translation techniques (CTDD and TDVR) present several advantages and drawbacks vis-à-vis TT.

Although TD in general has been under-investigated, apart from a few efforts in the years that preceded the massif influx of PCs (Gingold, 1978; Gouanvic, 1976; Laroque-Divirgilio, 1980, 1981) and a few recent pilot studies carried out with students and professionals (Dragsted & Hansen, 2009; Dragsted et al., 2009; Hétu, 2012; Mees et al., 2013), studies consistently show a better performance when speaking a translation than when typing it. For example, recent studies (Dragsted & Hansen, 2009; Dragsted et al., 2009; Mees et al., 2013) observe that TD is in general more efficient than TT since, instead of typing, the translator produces the translation orally – at the speed of speech – and focuses all his or her attention on the source-text copy. Nonetheless, as mentioned earlier in this chapter, one of the drawbacks specific to CTDD is the need to hire additional staff to carry out the transcription of dictations, when the translator or the client is unable to do so, thus leading to extra production time and costs. This being said, in some contexts where translators and transcriptionists still collaborate (for example one transcriptionist collaborating with three translators), CTDD is more efficient and, consequently, more cost-effective than TT; however, it is increasingly difficult to find skilled transcriptionists in the workforce (in Canada) (Y. Martin, personal communication, September 4, 2013).

TDVR provides the possibility of removing the transcriptionist from the equation, but has its drawbacks too. Unlike CTDD, real-time TDVR is considered to be "'mentally taxing' and 'more intensive, and requires a higher degree of concentration and attentional effort'" (Dragsted et al., 2009, p. 305). The reason for this is that real-time TDVR requires, or rather provokes, the translator to constantly shift attention from the source text to watch the target text transcription by the VR system as it appears on the screen, and to correct any transcription errors on-the-go in order to prevent errors from propagating (Dragsted et al., 2009, p. 313). This behaviour is to be further investigated, since it is possible for the translator not to focus his or her attention on the real-time transcription by the system (i.e., to ignore it completely), and rather do any revisions at the end of the whole translation process.

Offline TDVR is believed to mitigate or to eliminate those issues (Wu, 2008),[46] since the translator dictates first into a digital dictaphone and has the full dictation audio file transcribed later by a VR system, which processes the dictation as a whole instead of on a sentence-by-sentence basis. But for that approach to work properly, the translator has to purchase several devices, that is, a powerful PC and a good-quality digital dictaphone (pluggable to the computer via a USB port for instance). The translator also needs a good VR system with an audio file transcription feature, in addition to having excellent dictation skills (to dictate a text from top to bottom). The translator also needs to be able to dictate in environments appropriate for dictation (e.g. noise-free) and that are acoustically similar to where the VR system was

[46] *"The Most Cost-effective Way of Translation—Combined Use of Digital Recorders and Dragon Voice Recognition System".* Slides of a presentation at the United Nations in Nairobi on August 4, 2008. No publication followed the presentation.

originally trained, in order to ensure good word accuracy rates (in other words, a flawless or nearly flawless transcription).

In sum, despite the documented advantages of CTDD and TDVR vis-à-vis TT, their drawbacks and challenges remain significant. This is why it is appropriate and timely to explore ITD, because it introduces a ground-breaking scenario that may overcome the limitations of other types of TD: the combined use of VR with other emerging input modes in a multimodal environment (see section 4.3) which is expected to provide a better TX than the traditional KM workstation, and to perform better than VR in speech-only (monomodal) settings, as described above.

In chapters 1 and 2, I argued that the expansion of cloud computing, wireless technologies and voice-and-touch-enabled MIs have led humans to new ways of interacting with machines and of accessing enormous amounts of information and services from practically anywhere. In the present chapter, I explored the various translation techniques used throughout history and nowadays; I also discussed the extent to which oral techniques and VR technologies have been used and investigated in translation, highlighting their current potential and challenges. These challenges include developing the users' ability to combine different input modes (i.e., to interact multimodally), as well as optimizing VR systems and MIs to carry out translation-specific tasks, such as issuing (bilingual) voice commands for information retrieval, and enabling interaction in mobile contexts.

The following chapter will explore the literature on bilingual VR (BVR) and multimodal interaction, and will point out to the advantages and challenges of these emerging forms of HCI and HII. Subsequently, a description of ITD will be provided.

Chapter 4 — Beyond Unilingual Stand-Alone Voice Recognition Systems: Bilingual Voice Recognition and Multimodal Interaction for Translation

"Simplicity is the ultimate sophistication."
—Leonardo da Vinci

The present chapter examines bilingual VR (BVR) and multimodal interaction from a translation perspective. It presents a number of relevant works in the two fields and shows how these emerging interaction modes go beyond unilingual stand-alone VR systems and WIMP interfaces, currently used by certain translators for TDVR. It argues why these new approaches will need to be considered in future TCI and TII research. In light of this, a description of ITD will be offered, while highlighting the potential of ITD to become one of the most effective, efficient and ergonomic *modi operandi* in translation practice.

4.1. Bilingual Voice Recognition for Translation

While multimodal interaction is at the core of the present study and of the main experiment, it is worth exploring and discussing another potential approach to adapting VR technology to translation tasks: by enabling bilingual voice input. Indeed, the inability of commercially available VR programs to process two languages simultaneously is one of the technological challenges identified in my previous study (Zapata Rojas, 2012). While the central task of a translator, within an (I)TD context, is to dictate the target text (i.e., the translation) in the target language, he or she also needs to interact with the machine to carry out other translation tasks and make use of an array of tools, and this *may* be done bilingually (i.e., in both the source and target languages). For instance, to search for a term equivalent in a term bank, an English-to-French translator might want to launch a query such as: "*Rechercher l'équivalent en français de* bionic plant *dans ma base terminologique*". This would be a bilingual voice command.

Could a VR system process that? Let us explore this matter in detail. This exploration is also meant to support the appropriateness of multimodal interaction as a means to go beyond monomodal, stand-alone VR systems and to achieve a more natural interaction; to achieve simplicity. A discussion on this matter is also provided later in this chapter.

4.1.1. What is Bilingual Voice Recognition?

Let us first state what a bilingual voice recognition system (BVRS) is not. It is *not* an application that provides (bilingual) users with the possibility of selecting their language of choice before interacting with a machine through speech. Off-the-shelf VR systems with that feature have been mislabelled as "bilingual," and yet they simply integrate two unilingual systems that function separately. The user needs to deactivate the English system in order to activate the French system, for instance.

BVRSs, as I define them, are computer programs that allow voice interaction with a machine in two languages indistinctively and concurrently, either for issuing commands or for transcribing speech, without the need for the user to switch between two unilingual systems. To the best of my knowledge, full-fledged BVRSs cannot be purchased from software stores;[47] to date, a few efforts have been made to develop and implement such systems in a number of languages and for certain specific domains. Some of these efforts are outlined in the sub-section below.

4.1.2. Past Work on Bilingual Voice Recognition

Outside of translation research, some breakthrough experiments with BVRS have taken place over the past 15 years in bilingual contexts, that is, in societies where two (or more) languages coexist and/or have the same status, and in some domains where bilingual—and even multilingual (Uebler, 2011)—voice input is desirable, even though those languages do not coexist.

An example of a country where several languages coexist—and share the same status in certain regions—is Spain. Several efforts to develop BVRSs have been made throughout the Iberian Peninsula. For instance, Mariño et al. (2000) carried out an experiment with unilingual Spanish and Catalan VR systems, and a bilingual system for the two languages, developed in their laboratories in Barcelona. Spanish and Catalan are distinct enough so that unilingual speakers of one language or the other cannot readily understand the other. However, since both languages are derived from Latin and have long coexisted in Eastern Spain, their phonetic repertoires (or phone sets) have important similarities. Such similarities facilitated the development of a Spanish-Catalan BVRS. In this experiment, unilingual and bilingual systems were built from large spoken corpora by creating acoustic and language models in both languages. To test the performance of the unilingual and bilingual systems, two tasks

[47] The recently unveiled Google Voice Search app for Android devices is an exception. See discussion below.

were chosen. The first was an isolated-word test (utterances of people's names and cities); the second involved uttering telephone numbers. It is important to note that the Spanish and Catalan unilingual VR systems were designed with a reduced set of allophones[48] in relation to the full inventory of allophones of each language. As for the BVRS, tests were carried out using a clustering algorithm in order to confirm the acoustic similarity between the Spanish and Catalan sounds transcribed with the same allophone, and to assess whether certain allophones could be grouped or not. The BVRS was developed from a corpus composed of a comparable amount of speech for both languages; therefore, a balanced training of Catalan and Spanish was achieved. Its performance was observed to be equivalent to that provided by the Spanish system developed using speech from Eastern Spanish (spoken in the same part of the country where Catalan is spoken) and by the Catalan system. By the end of their experiment, Mariño and his collaborators concluded that the Catalan-Spanish BVRS saves parameters without a loss in performance with respect to the unilingual systems.

Not far from Catalonia, Alabau and Martínez (2006) performed an experiment for the Valencian language, a dialect of the Catalan language spoken in the Spanish *Comunitat Valenciana*. This dialect, however, is different from other Catalan dialects in that its phone set differs very little from the Spanish language. For this experiment, the corpus used to create the acoustic models for the BVRS was considerably smaller than for the Catalan-Spanish BVRS experiment described above (one hour of speech,

[48] An allophone is one of a set of possible realizations of a speech sound (or phone) used to pronounce a single phoneme.

compared to about 30 hours of speech per language). The spoken corpus contained a two-hour recording of comparable utterances by native speakers of each language, male and female, thus ensuring a balanced distribution of utterances. In the experiment, good results were achieved with the Valencian-Spanish BVRS, most likely because with such similarity in phone sets, more training data could be used for the same model. The researchers pointed out, however, that for languages that differ considerably, the performance might not be as good as it was in this experiment.

Galician is another language spoken in Spain, in the Galician Region located in the north-western part of the country. Diéguez-Tirado et al. (2005) carried out an experiment to test a Galician-Spanish BVRS specifically intended for transcribing news programs on TV. The spoken corpus used to build the language models was taken from broadcasting news recordings. Now, given the changing acoustic conditions of broadcasting, the variety of speakers, styles and topics possible, not to mention the added problem of switching between languages (or code-switching), this experiment was rather complex. Diéguez-Tirado and his collaborators wanted to avoid a priori language detection, and designed a system that could take into account both languages simultaneously. Again, comparable to the Catalan-Spanish and Valencian-Spanish experiments outlined above, this is possible because of the phonetic similarity of Galician and Spanish. However, the results of this particular experiment showed an unbalanced performance in Galician and Spanish because the Galician portions of the corpus corresponded to planned speech, whereas the Spanish portions contained large stretches of spontaneous speech (e.g. people being interviewed on the street), leading to varying speed and language incongruences, which made it more difficult to train the BVRS.

Quite far from Europe, India offers another example of coexistent multilingualism. Several dozens of languages are spoken across the country and often coexist with other official languages. Thus, bilingual and even multilingual VR systems would be desirable in the case of India. Kumar and Say Wei (2003) carried out a preliminary experiment with two phonetically and syntactically different languages (Tamil and American-accented English) in order to test the viability of BVRSs in India. Having carried out their experiments, Kumar and Say Wei observed that, in the BVRS, the English and Tamil acoustic models for each language (trained separately) became general enough to cater for both languages.

Likewise, other experiments have been carried out with languages that differ considerably, even though the languages do not coexist. Examples of this are Yu et al.'s (2003) and Zhang et al.´s (2008) experiments for English-and-Chinese BVR. Likewise, Žibert et al. (2003) proposed a Slovenian-Croatian BVRS to be integrated in a spoken-dialogue system for weather forecasts. This system was aimed at initially identifying the user’s language and then providing the user with weather-related information in that language. No inter-sentential or intra-sentential switching was expected.

Lastly for this overview, let us mention that the Google Voice Search application for Android devices has recently enabled multilingual functionality,[49] but such an improvement is limited to search queries and simple commands to the system. Users can launch simple queries on Google or commands to the device's operating system in up to five preset languages, with no intra-sentential code-switching; only one query or command in one language at a time.

This brief literature review clearly shows a remarkable interest in developing VR systems that can process two or more languages, without the need for the user to switch the languages or, in other words, deactivate one language and activate another. These initial efforts shed light on the various contexts and domains where bilingual voice interaction with computers is desirable and viable. Such interaction is thus not limited to obtaining speech transcription in two languages; it is also desirable and viable for issuing bilingual commands to computer systems, for instance, to retrieve information. The Google Voice Search app for Android devices is a first successful effort towards enabling VR in two languages or more.

In the following section, I provide some insight from the point of view of translation as a potential scope for BVRSs.

4.1.3. Discussion

As it has been argued so far in the present thesis, there is strong evidence of the need to open up new perspectives in TS research to design and develop tools that can meet

[49] Source: http://www.cnet.com/news/que-pasa-google-now-adds-true-multi-language-support/. Retrieved on 23-11-2015.

the changing professional needs of translators and cope up with the evolution of technologies. The focus is expected to shift towards VR because it seems to open up new perspectives that other NLP applications have not done so far.

As stated at the beginning of this section, while the central task of a translator in a dictation context is to dictate the translation in the target language, he or she also needs to interact with the machine to perform other translation tasks, and this *may* be done bilingually. The hypothetical example provided is that searching for a term equivalent in a term bank using one's voice would entail launching a command such as: "*Rechercher l'équivalent en français de* bionic plant *dans ma base terminologique*".

Voice input is desirable not only for transcribing speech, but also for issuing voice commands to the operating system and to the various applications used by translators, who constantly switch between languages while working. Now, in light of the experiments outlined above, should BVR be considered when designing and developing new computer tools for translators, and in particular an ITD tool?

The reader may agree that the literature review above, although far from comprehensive and representative of every language combination and bilingual context, is probably enough to affirm that BVR is indeed possible. However, it is not enough to determine whether bilingual voice input is absolutely necessary for translation-related tasks; to the best of my knowledge, no experiment has investigated BVR in translation.

Voice input is just one of the several input modes currently explored in HCI and HII as alternatives to traditional KM input. While possible, an empirical experiment investigating BVR for translation purposes is beyond the scope of the present doctoral work, but remains to be explored in future investigations.

Furthermore, as will be discussed in the following section, and as suggested already in the introductory chapter, the focus of this work is on multimodal interaction because past experiments in this field have proven that voice input, when combined with other modes such as touch and stylus, turns out to perform better than in voice-only contexts; bilingual or not, may I add. For instance, in the case of the hypothetical bilingual command "*Rechercher l'équivalent en français de* bionic plant *dans ma base terminologique*", the English portion of the command could be obviated by combining another input signal (see discussion later in footnote 50 on page 81); there would be no need for a system that processes VR in both English and French simultaneously. But BVR is not to be ruled out completely, as it may even complement multimodality, adding a new dimension to the workflow and augmenting the users' interaction possibilities. Such exploration will be on hold for future work.

Now, let us now explore multimodal interaction in further detail by examining some of the literature in the field.

4.2. Multimodal Interaction for Translation

Multimodal interaction represents a more efficient, flexible and expressive means of interaction with machines; than KM-based interfaces; a paradigm shift away from traditional WIMP interfaces. One of the goals of multimodality is to recognize natural forms of human behaviour and language in search of a more natural interaction with machines (Oviatt, 2012, p. 405). In this section, based on past work on multimodal interaction, I will validate why studies in this field, such as the one reported on in this thesis, are timely and relevant in TS research.

4.2.1. Past Work on Multimodal Interaction

The first efforts to develop and test multimodal systems took place during the 1980s and 1990s. These works coincided with the ongoing efforts for the improvement of VR technology and its integration into different professional domains and applications (Oviatt, 2012, p. 406; Rabiner, 1997). One of the earliest works in multimodality is probably Bolt's famous "Put that there" experiment (1980), which examines some of the advantages of combining different input modes such as voice and pointing. In his seminal paper, Bolt describes an experiment in which a person issues voice commands to a computer system and combines 2D pointing gestures on an armrest-mounted touchpad in order to carry out different tasks such as creating, modifying and moving objects. Some of the systems developed in the years that followed Bolt's experiment combined voice with other input modes. For example, Cohen et al. (1989), based on the use of two conversational agents (computer systems with which a user can have a conversation) (Weizenbaum, 1966), highlight the advantages of voice interaction while lending support to the idea of combining several input modes to hinder the problems and limitations of monomodal interaction, that is, the use of one input mode at a time. Likewise, Wahlster (1991) explores deictic gestures by analyzing gestural and linguistic communication between humans and by comparing it with the interaction of humans with machines. Likewise, Neal and Shapiro (1991) describe the prototype of the CUBRICON system, a MI that combines different input and output modes such as voice, keyboards, gestures, tables, shapes, maps and images.

Subsequent works explored systems that combined voice, touch and stylus input, or VTS systems. Although this combination is not necessarily *the* most efficient approach in multimodality, VTS systems have advanced more rapidly in their architectures than MIs that combine voice and gaze, or voice and head or full-body movement, for example (Oviatt, 2012). For this reason, a stronger focus is put on VTS systems for the purposes of this study. Another reason for this is that translation is primarily an act of written and oral communication; thus, other interaction modes may result rather peripheral, yet possible and viable. Recent studies have shown more and more robust architectures of systems involving, for instance, voice and gaze (Misu, 2015), and gaze and free-space hand gestures (Chatterjee, Xiao, & Harrison, 2015).

But what are the advantages of multimodal interaction? In the following subsection, I will present some works carried out over the past decades in an effort to answer that question.

4.2.2. Advantages of Multimodal Interaction

A fair number of advantages are attributed to multimodality, both from UX and computational viewpoints.

Among the documented advantages from the human perspective are efficiency gains. Cohen et al. (1998) describe a comparative study of the interaction with a GUI and with a VTS interface, and report efficiency rates three to nine times higher with the

VTS interface. These efficiency gains are attributed to the system's ability to process several input modes in parallel.

Furthermore, Oviatt and Cohen (2000) argue that MIs lead to new ways of using technology, for instance, in mobile contexts or outdoors. Such interfaces would allow the user to alternate, if need be, between the different input modes according to the changing conditions of mobile/outdoor interaction. Likewise, VTS can be considered as active input modes, since they are clear indicators of the user's intentions (Bangalore & Johnston, 2009; Oviatt, 2012, p. 408), as opposed to passive input modes such as the mouse (see footnote 29 on page 42). Indeed, in a VTS system, a natural language is processed in parallel with complex gestures performed with the fingers or stylus on the touch-screen, leading thus to hundreds of symbolic interpretations, since an individual may use his or her fingers or a stylus not only for pointing, but also for making annotations or drawings (Oviatt, Cohen, Fong, & Frank, 1992; Oviatt, DeAngeli, & Kuhn, 1997; Oviatt, 2012, p. 406).

Another advantage of MIs from the human perspective is the flexibility to use different input and output modes: the user may combine different modes or alternate between them at any time (Bangalore & Johnston, 2009). Some comparative studies show the users' level of frustration when working with a monomodal system and how that frustration is reduced when they have the choice to alternate between various input modes, since errors can be easily avoided or repaired. For example, Oviatt and van Gent (1996) observe that users select the input mode they believe is less likely to produce errors in a particular situation. In other words, multimodal interaction would allow error avoidance.

In addition, MIs would have the ability to reduce the user's cognitive load. Users have shown a tendency to go from monomodal to multimodal interaction as the complexity of tasks increases (Oviatt, Coulston, & Lunsford, 2004).

Besides, MIs are considered to be more efficient from the design and development viewpoint. One of the advantages of multimodal systems that combine VR and other input modes is that NLP is less challenging from a computational perspective, since more direct commands can be issued to the system when interacting multimodally, thus the system has a better ability to process the various input signals. This is what is called mutual disambiguation of input signals (Oviatt & Kuhn, 1998; Oviatt, 1999). For instance, a combination of pointing and a few spoken words can be used to issue a command that would normally need, in a monomodal VR environment, many more words—and much more computational power (Oviatt & Kuhn, 1998).[50]

Furthermore, UI designers often have to face the challenges of designing for multiple users, who may have different profiles. In fact, the great diversity between human beings (races, cultures, languages, abilities and handicaps, among other factors)

[50] Another example is Bolt's 1980 experiment described above. Three words (i.e., "put that there") and pointing gestures will suffice to replace a longer monomodal voice command and such as "*put the smaller blue circle at the top right corner of the screen.*" Consider also the example of the hypothetical bilingual command from the discussion on BVR above: "*Rechercher l'équivalent en français de* bionic plant *dans ma base terminologique*". A combination of input signals (e.g. voice and on-screen pointing), *and* context-awareness by the system, could reduce such a long command to something like "*rechercher ça*" (search that), issuing the command after selecting with one's finger or a stylus the character string to be searched (in this case, "bionic plant").

reminds us that there is no image of an average user. Now, MIs allow different groups of users to determine the way they interact with the system. Not only do MIs have the potential to function more effectively and efficiently than monomodal interfaces, they also do it in a more stable way in different contexts and for users with different profiles. Behavioural differences have indeed inspired a number of studies in multimodality. Oviatt et al. (1994), for instance, point out that linguistic variability is problematic when designing MIs that process natural language. In the light of the results of several empirical studies examining the way different individuals speak and write, these researchers confirm the need to design MIs that minimize the problems attached to the user's linguistic variability. Another study (Oviatt, Lunsford, & Coulston, 2005) examined the various forms of multimodal interaction (i.e., whether multimodal interaction is done simultaneously or sequentially) and if participants, who belonged to different age groups, showed other behavioural characteristics that may offer a coherent explanation to their fundamental differences, in order to guide the design and implementation of new interfaces.

Developments in multimodality have gone beyond research-level systems. In our day, among the most concrete and widely known-and-used examples of commercial MIs are tablets and smartphones, which are basically voice-and-touch-enabled. Schalkwick et al. (2010) describe a smartphone interaction case study in which individuals use voice input to perform web searches and to issue commands to the phone's operating system. While acknowledging the challenges of designing and developing MIs, these researchers highlight the many advantages attributed to this type of interaction. Hence, smartphones clearly exemplify the advantages of mobility and multimodality: an increased efficiency and flexibility; in other words, a more natural interaction with the technology.

Let us note that multimodal interaction has been gaining ground in recent TS research (Alabau & Leiva, 2014; Garcia-Martinez et al., 2014; Mesa-Lao, 2014). In Mesa-Lao's study (2014), for instance, student translators surveyed reported that they would welcome the integration of VR as one of the possible input modes for performing translation tasks. Eighty percent of respondents (n=15) stated that they would definitely consider adopting VR in combination with no-speech input modes in their daily practice as translators and, in the case of the study, for post-editing MT output. VR is thus a complement, rather than a substitute, for non-speech input methods.

In the same vein, as a result of the popularization of touch-and-pen-enabled multimodal devices, the potential for simulating the way people in general (and translators in particular) use pen and paper,[51] is explored. Oviatt (2012, p. 411; 2013) notes that pen input supports better problem solving in science, technology, engineering and mathematics than existing keyboard-based GUIs. Advantages attributed to pen interaction include improved efficiency and the ability to focus attention, to solve problems correctly, to engage in high-level synthetic thinking, to communicate fluently and to memorize information. Oviatt also argues (2012, p. 412) that pen input is:

[51] MyScript Notes Mobile is an example of a pen-and-paper-interaction-like application for mobile devices, which relies on spatial relations of the ink strokes to perform handwriting recognition. This type of system, however, does not seem suitable for translation purposes (Alabau and Leiva, 2014).

> a rich communications modality that is capable of expressing different representations (linguistic, symbolic, diagrammatic, numeric), can facilitate people's communicative fluency, ideational fluency, and problem solving to a greater extent than either keyboard-based [GUIs] or non-digital pen and pencil.

As a matter of fact, as mentioned in chapter 2, pen input is also explored in TCI, for instance, for the purposes of proofreading translated texts (Alabau & Leiva, 2014)[52]. In their work, Alabau and Leiva (2014) propose a taxonomy of proofreading gestures (word change, letter case, punctuation, word combination, selection and text displacement gestures) (see Figure 6 below) which may be employed in a multimodal translation proofreading interface.

[52] Other experiments in interactive text proofreading include Coleman (1969), Leiva et al., (2013), Shilman et al. (2006) and Romero et al. (2009).

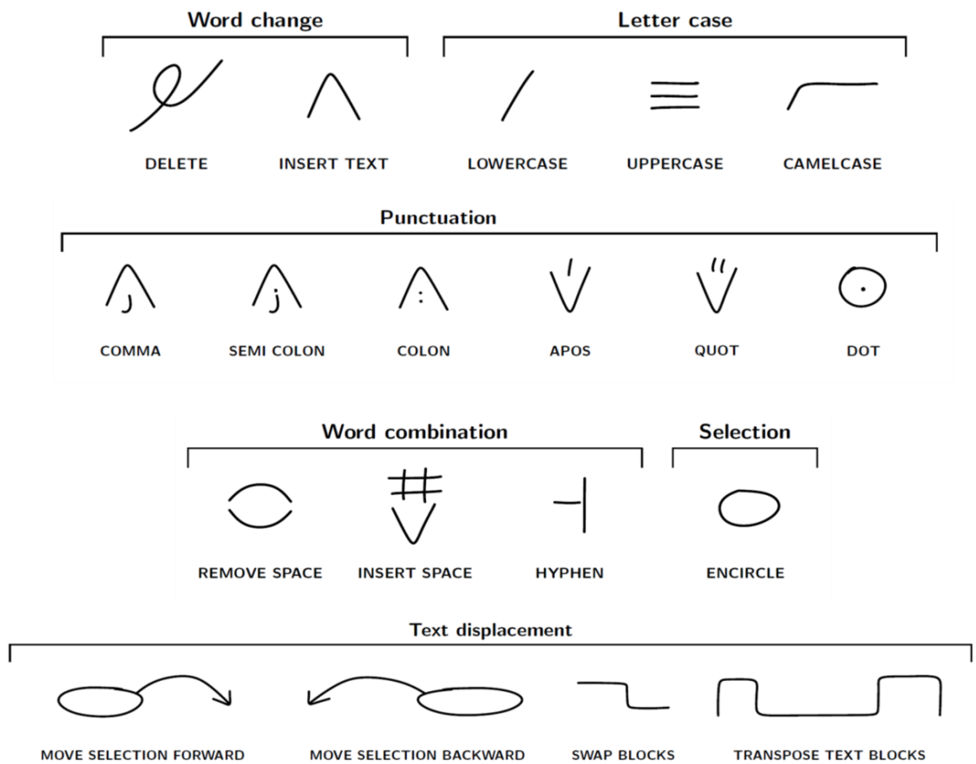


Figure 6. Taxonomy of handwritten gestures for translation proofreading purposes. In Alabau and Leiva (2014).

Lastly, regarding multimodality in translation, O'Brien et al. (2014) introduced the first prototype of a smartphone-based CAT tool, specifically for post-editing purposes. The researchers are considering the possibility to integrate voice input to the application, in addition to the current touch-and-swipe capabilities. The tool is in its earliy stages and TX data is currently being collected in an effort to improve it. It is worth noting that the initial user feedback was positive.

The many possibilities that multimodality may provide to translation practice are still to be unveiled. This literature review reveals that multimodal interaction is a promising direction in the search for translation tools (and technology applications in general) that provide users with a better UX and that take the human factor into account. Thus, a multidisciplinary perspective will be more central now than it has been traditionally for GUI design (Oviatt, 2012, p. 412) and, more specifically, for translation tool and environment design and implementation.

The present study on ITD is an excellent start towards generating definite change and innovation in TCI and TII research. But how is ITD different from other translation techniques and other approaches to CAT? Let us now take a closer look at ITD.

4.3. Interactive Translation Dictation: How Is It Different?

ITD has the potential to become one of the most effective, efficient and ergonomic *modi operandi* in translation practice. Let us recall that ITD, as defined in this thesis, is a translation technique that involves interaction with voice-and-touch multimodal interfaces, throughout the entire translation process, namely during preparation, production and revision. Although the notion of ITD is somewhat raw, it already suggests a different form of TD (see Figure 3 on page 62): the possibility of using one's voice in combination with other input modes on different computational devices to carry out the bulk of translation tasks.

In ITD, there are different ways of using VR and other emerging input modes such as touch and stylus input throughout the translation workflow. First, during the preparation step, voice commands are issued to the system in order to retrieve the information (e.g. term equivalents, encyclopedic information) needed to produce the translation. Let us point out that, in many UIs, commands have proven to be important determinants of rapid learning and use, low error rates, high satisfaction and easy retention over time (Shneiderman, 1988, p. 703). Shneiderman (1998, p. 279) explains that "the essence of command languages is that they have an ephemeral nature and that they produce an immediate result on some objects of interest". Thus, UIs that combine command language and natural language (e.g. UIs that respond to voice commands) have long been interesting for developers for they allow users to use a natural language to give instructions to the computer and receive (almost-immediate) responses from it (Shneiderman, 1998, p. 294).

In the ITD scenario I propose, during the production step, the translation is dictated into the device's VR system, a similar approach to TDVR. Lastly, revisions (additions, deletions and modifications to the text) and formatting can also be done through voice commands. Moreover, stylus and/or touch can be combined with voice input at any step of the process for pointing or for text input using a virtual (software) keyboard[53], as illustrated in

Figure 7 below:

[53] Commercial MIs such as tablets and smartphones are usually equipped with a virtual or software keyboard, or even "ultra-fast" virtual keyboards that allow users to swipe their fingers on the keyboard instead of "pressing" each "key" on the screen. MIs' virtual keyboards also have automatic word correction and word suggestion features, and sometimes the option to personalize the user's dictionary (i.e., the recognizable set of words) in order to improve recognition and accuracy.

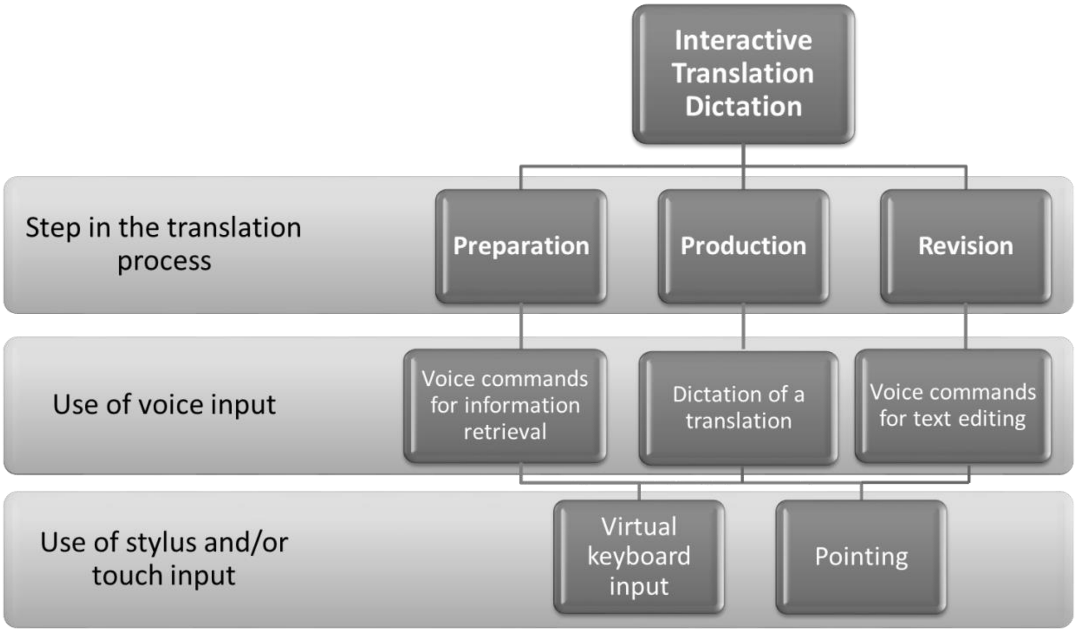


Figure 7. Interactive translation dictation. Uses of voice-recognition, stylus and/or touch at the different steps of the translation process.

Up to this point, I have elaborated on the demonstrated success of VR technology and oral translation techniques, and on the potential of combining input modes such as VTS, as well as the advantages of multimodal interaction both for users and developers. This doctoral study conceives and tests new ways of using these techniques and emerging technologies in translation to provide a basis for better-grounded TCI and TII research and, more specifically, for the design and development of a multimodal ITD environment. Multimodality is, in my opinion, one of the best possible ways to circumvent the current shortcomings and limitations of existing CAT tools, of KM interfaces, and of stand-alone (monomodal) unilingual VR systems.

The next two chapters provide a full description of the experimental framework of this study, and present and discuss the quantitative and qualitative data collected.

Chapter 5 — Translating with Multimodal Interfaces: An Experiment

"A thinker sees his own actions as experiments and questions as attempts to find out something. Success and failure are for him answers above all." — Friedrich Nietzsche

This chapter provides a full description of the experimental framework of my doctoral study. First, I explore the challenges of measuring usability and elaborate on the choice of measures to report on the TX with a multimodal ITD environment vs. a KM environment. Then some data-collection approaches commonly used in translation process research (TPR) will be discussed. I will then elaborate on the experimental setup and methodology.

5.1. The Challenges of Measuring Usability and Choice of Measures for the Study

Measuring the quality and performance of different computer tools or interfaces has gone on for a number of decades now, first under the headings of human factors or ergonomics, and later of usability and UX (see section 1.3.1). The question of what to measure and, therefore, which data-collection methods and tools to choose, has been central to the design and development of new UIs and the improvement of existing ones.

Investigating the usability of computer systems and devices has long been a key element of HCI research. However, as supported by Hornbæk (2006), usability cannot be directly measured. Instead, it is possible to investigate the quality-in-use[54] of a system or interactive product, which involves observing and analyzing the UX with the system or product (Bargas-Avila & Hornbæk, 2011). Now, although thoroughly investigating quality-in-use and UX has been considered a highly challenging task, it is nonetheless possible for researchers to operationalize the usability construct to find aspects that can actually be measured and analyzed.

As stated in the introductory chapter, usability measures have traditionally reported on three aspects of the interaction: (1) effectiveness (measures of the quality of or the satisfaction with the outcomes); (2) efficiency (measures of the interaction process); and/or (3) user satisfaction (measures of users' attitudes and experience). This doctoral study uses some of the measures gathered by Hornbæk (2006) (see Table 1 below), and attempts to demonstrate how the various aspects of usability, as well as subjective and objective measures, may have (unexpected) relations. Considering the scope and limitations of the study, I chose the measures that were more likely to respond to the research question. This choice was made based on pilot experiments and discussions with field experts. Table 1 was amended during the data-analysis and thesis-writing process as I discovered that some usability measures did not align with the goals of this study properly. The analysis looks at objective and subjective measures separately, but provides insight into the possible relations that may be found between them.

[54] As a reminder to the reader, Bevan (1995) defines *quality-in-use* as "the extent to which a product satisfies stated and implied needs when used under stated conditions."

Indeed, the combined analysis of subjective and objective measures is a central aspect of this study. Subjective measures of usability are measures that "concern users' perception of or attitudes towards the interface, the interaction, or the outcome" (*ibid*., 91), whereas objective measures are not dependent on users' perception. The following are the sets of objective and subjective measures reported in this thesis.

Usability Aspects	Objective Measures	Subjective Measures
Effectiveness (Outcomes)	**Expert assessment of the outcome, error rates**	**Users' perception and assessment of the outcome**
Efficiency (Interaction Process)	**Task completion times, time in mode, usage patterns, input rates, deviation from an optimal solution, other**	**Subjectively experienced duration, users' perception of interaction and task difficulty**
Satisfaction (Users' Attitudes and Experiences)		**context dependency, ease-of-use, users' sense of success, other**

Table 1. Choice of objective and subjective measures per aspect of usability.

The distinction between objective and subjective measures has been extensively dealt with in the literature. According to Hornbæk (2006), one of the reasons why it is necessary to consider both subjective and objective measures is that they may lead to different conclusions regarding the usability of an interface. For instance, examples have proposed as a usability measure the ratio between objective time and subjectively experienced duration. Psychologists have long recognized and quantified the difference between these two measures; therefore, using both may lead to a better picture of usability and of the UX, since "differences between interfaces in objective time may not be found for subjectively experienced duration, and vice versa" (*ibid*., 92). Likewise, the importance of task times as a measure of usability may have been overestimated in a number of studies since users, designers and owners of an interface may not equally weigh their importance (*ibid*., 80). Indeed, higher task completion times have been observed to be indicators of users' motivation, reflection, and engagement (Inkpen, 2001).

The lack of longitudinal studies has been considered a major challenge in usability research. Long-term studies are actually rare and expensive. Little is known about how usability develops as the user spends more and more time interacting with the interface, and how trade-offs and relations between usability aspects change over time, and if users are able to eventually compensate for most problems and difficulties that lead to dissatisfaction when interacting with the interface for the first time (Hornbæk, 2006). Setting up a longitudinal study is difficult due to a number of factors inherent in empirical research (see, for instance, the limitations for the present study in section 1.6).

Lastly, the understanding of the UX and the usability of an interface also depends on whether tasks are considered at a micro level or a macro level (*ibid*.). On the one hand, at the micro level, measures cover tasks of short duration, that have a manageable complexity, and that focus on perceptual or motor aspects. At this level, task times are usually a critical aspect of the interaction. On the other hand, at the macro level, measures cover tasks that are longer (e.g. that last hours, days or months), that are cognitively or socially complex, and that display large individual differences in the interaction process and vast variations in the outcome. At this level, effectiveness and satisfaction are critical parameters (*ibid*.).

In this study, the working model proposed by Hornbæk in his 2006 paper could not be adopted in full due to a number of limitations and challenges inherent in research in the humanities and to research ethics (in particular related to anonymity). However, as mentioned above, a first attempt to exploring the possible relationships between the various measures and aspects of usability will be made here.

5.2. Selecting a Data-Collection Approach

The experiment was carried out with the following research question in mind: Does the physical prototype of an ITD environment consistently provide translators with a better TX[55] than a traditional KM environment, when preparing, producing and revising a translation? The hypothesis is that the prototypical ITD environment consistently provides translators with a better TX than the KM environment, considering the translation process as a whole.

Studying and improving the UX and the quality-in-use of computer systems and tool prototypes has been a key element of HCI (and will constitute, without a doubt, a major focus of all research in TCI and TII). It is thus essential to include real users working in real-life (or simulated, close to real-life) situations when investigating the

[55] A reminder to the reader that the notion of TX is introduced in this research. Based on ISO's definition of UX (2010), TX is defined as a translator's perceptions of and responses to the use or anticipated use of a product, system or service.

UX and the performance of new tools and prototypes, which in return helps designers and developers in making key decisions about particular aspects and features of a UI.

As a matter of fact, human-centred research is far from new in TS research. Several methods have been used for a number of years to study human translator behaviour. TPR scholars have long investigated not only the cognitive processes involved in translation activity, but also the way translators interact with computers and with computer tools. In the subsection below, I will present some of the data-collection approaches traditionally used in TPR. In the subsequent section, I will elaborate on the methods and tools selected for investigating the TX, and on why I favoured those methods and tools over other approaches within the framework of this study.

5.2.1. An Overview of Data-Collection Approaches in Translation Process Research and Choice of Approach for this Study

In TPR, different data-collection methods and tools have been used to learn about how translators translate, to speculate about what happens in the mind of a translator during the translation process, and to observe how they interact with various computer tools and devices (Teixeira, 2014). Traditional data-collection methods include mainly think-aloud protocols (TAPs), keyboard- and mouse-activity recording, and eye-tracking (using specialized TPR software).

In the earlier TPR years, TAPs (Ericsson & Simon, 1984) were the main method for investigating the cognitive information-processing involved in translation. The think-aloud approach consists of asking participants in a study to verbalize their thoughts as they are carrying out a given task. Verbalizations may be logged, recorded and analyzed in various ways.

A few years later, with the introduction of key-logging applications, it was possible to combine TAPs and keyboard activity, which allowed researchers to triangulate data, carry out deeper observations and analyses, and formulate stronger hypotheses about translation processes in general. Some researchers (Jakobsen, 2003; Teixeira, 2014) believe, however, that the think-aloud method adds to the cognitive stress the translator is working under. In addition, TAPs are also believed to have an impact on translation speed, since translators have to work in an unnatural situation, thus affecting the "ecological validity" of experiments (*ibid*.).

Some other approaches have also been used to log and analyze translators' activity. For instance, the Translog software, developed specially for TPR in the 1990s, was conceived as an alternative to TAPs, since it runs in the background and does not interfere in the translation process; it logs all keystrokes and mouse clicks performed by the translator. With Translog, it was also possible to study cognitive phenomena such as chunking, based on the distribution of pauses, and other activities such as edits and corrections (Jakobsen, 2011). However, with the earliest 1990s version of the Translog program, it was only possible to speculate on what comprehension processes a translator engaged in before a solution was typed.

With such a limitation in mind, eye-tracking technology was added to the experimental environment, allowing researchers to observe and analyze what happens (or more precisely, where the translator's eyes go) when no typing activity is recorded (Jakobsen, 2011). According to Jakobsen, eye-movement logging, combined with key-logging, helps to better understand the mechanics and complexity of the written translation task, and to penetrate deeper in the cognitive processes involved in translation. In other words, by visualizing the way translators read and how their reading interacts with their typing, it became possible to obtain a better picture of how translation processes are coordinated in the mind of the translator. One of the assumptions is that integrating eye-and-keyboard data provides new opportunities for researching the dynamics of how source text comprehension activity is coordinated with target-text production activity (*ibid.*). But there are a number of challenges inherent to eye-data collection and analysis (O'Brien, 2009; Teixeira, 2014), as I will discuss now.

Eye-trackers are not all created equal. Remote eye-trackers are generally the preferred type in TPR, since translators are physically unaware that their eyes are being tracked. However, during a translation task, translators can spend considerable time just looking at the keyboard or away from the computer screen; therefore, experimental data on eye movement may be noisy or incomplete. Eye-tracking glasses may be used instead. However, the glasses are believed to be more invasive than remote eye-trackers, since translators have to wear the equipment on their heads. That being said, the glasses allow tracking movements outside the screen area. Head-mounted eye-trackers are generally considered to be more accurate than remote eye trackers, but affect the ecological validity of experiments, since translations are done in a rather unnatural working situation (O'Brien, 2009; Teixeira, 2014).

Some other challenges of using eye-tracking technology in TPR include the (high) costs of the devices, rapid technological redundancy (i.e. that they rapidly need to be upgraded to newer computers and operating systems), the learning curve (the time needed to learn new feautres and functionalities) and, perhaps more importantly, the accommodation for the eye tracker. Indeed, since a normal work environment is an important factor for an experiment's ecological validity, the eye-tracker needs to be set up in such a way that the participant is familiar and comfortable with it (O'Brien, 2009), but this does not seem to be possible for researchers to guarantee.

Different data-collection approaches used in TPR have led to different challenges, and new methods of investigation have been proposed over the years, as the nature of the interaction of translators with computer devices and with information evolves (see Carl et al., 2016). One thing remains certain, the more the research community becomes multidisciplinary and embraces new methods of investigation, the more mature the research will become, and this will be to the advantage of all those who are interested in human-centred translation research (O'Brien, 2009).

The present study joins current efforts in human-centred research in TS. The aim of this study is to observe and analyze the TX with MIs. The experimental design needed to take into consideration the various limitations of a doctoral project (in the humanities) (see section 1.6) and the various challenges inherent to the research methods and tools commonly used in TPR described above.

The experiment presented in this thesis proposes an unprecedented TCI scenario and a ground-breaking data-collection approach that combines active observation of users with video recording (using two camera feeds), screen recording using Screencast-O-Matic,[56] input logging using Translog II (Carl, 2012a) and InputLog (Leijten & Van Waes, 2013), and semi-structured interviews.

In the following section, I will provide a detailed description of the experimental setup of the study, including a description of a pre-control task that was included in order to complement the TX data.

5.3. Methodology

5.3.1. Pre-control Task

The experiment was carried out with 14 professional translators.[57] All participants performed a pre-control task at the beginning of the experimental session. The inclusion of this pre-control task in the study was motivated by an informal pilot experiment I carried out a few months earlier, the results of which I discussed informally within the framework of a translation data analytics project conducted in Copenhagen, Denmark, in August 2014. The pilot experiment was an exploratory approach to defining “translator efficiency”. A discussion of the notions, of the methodology and the results of this experiment is beyond the scope of this thesis. However, the reader should know that the findings of such an exploration allowed me to formulate the following hypotheses (yet to be validated): Provided that VR technology today has reached almost perfect accuracy rates and that the transcription is done quasi-simultaneously to speech, a translator is efficient when he or she can

[56] For information on this tool, visit: http://screencast-o-matic.com/home.

[57] More details on the participants’ profile will be provided in the subsequent sections and in chapter 6.

work on the production frontier (Coelli et al., 2005, p. 3), or very close to it. The production frontier (i.e., the maximum output attainable, or optimal efficiency) in translation could be reached by translating at the speed of speech (speaking at a normal pace and in a clear manner, as if giving a speech or having a conversation with someone). As a consequence, a tool/environment is technically efficient when it allows the translator to work on the production frontier.

The pre-control task consisted of transcribing (retyping) a French text, and reading out loud the same text. The aim of this pre-control task was to collect data on the participants' typing speed, transcription accuracy, reading-aloud times and other input data such as mouse activity. The text used for this task consisted of one title and one paragraph for a total of 130 words (or, more exactly, 862 characters, including spaces). The text used in this pre-control task is attached in appendix B.

The Translog II software was used during the two sub-tasks (i.e., typing the text and reading it out loud). Although Translog II was primarily designed to investigate human translation processes, it can also be used to study reading and writing processes in general (Carl, 2012b), as it was the case of this pre-control task.

The two sub-tasks were performed separately. First, participants were asked to launch the user's application (i.e., *Translog II User*) and open the project file for this experiment. The project had been set up (in *Translog II Supervisor*) to display the model text at the top of the window, so that participants could retype the text on the text field at the bottom half of the screen, as illustrated in Figure 8 below:

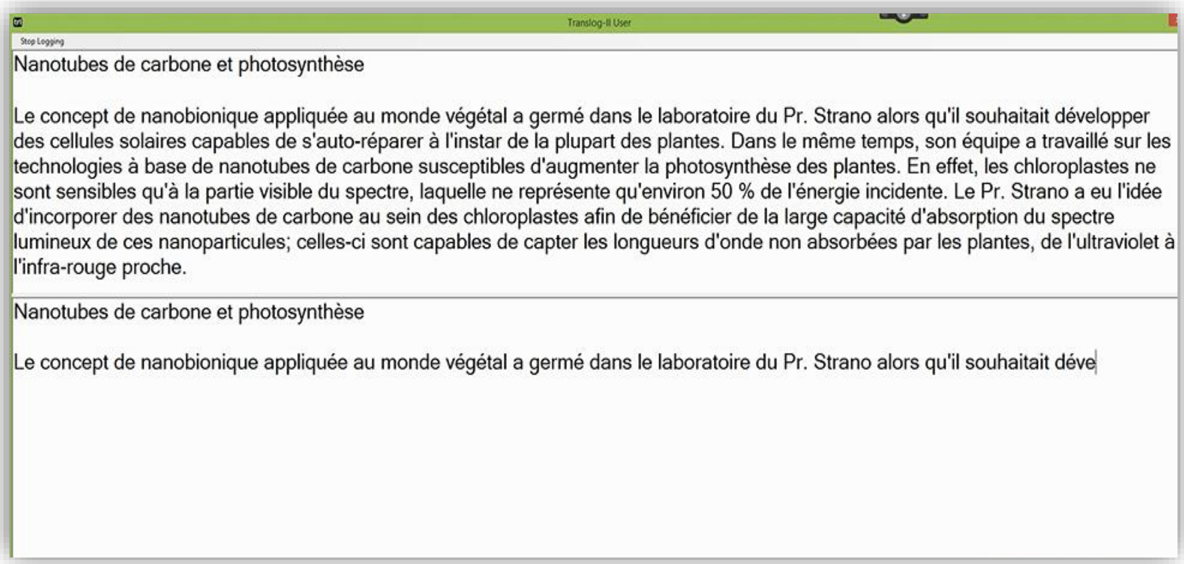


Figure 8. Translog II's User interface.

Once the copying task was completed, the logging session was saved, and a new session with the same project file was opened, this time for the reading-aloud sub-task. Participants were instructed to read aloud the same text at a normal pace, not too fast, not too slow, as naturally as possible, as if they were "broadcasting the news on TV or radio". Since screen/audio recordings were running in the background,[58] this session did not have to be saved in Translog II, because the analysis could be done with the recordings at a later time.[59]

5.3.2. Main Experiment

The main experiment consisted of comparing the TX in ITD versus KM interaction.

[58] More details on the experimental setup are provided below.

For that, I simulated an ITD environment. Simulations play a critical role in UI design: design sketches, physical prototypes, low-fidelity mock-ups and high-fidelity (automatic) simulations help to plan the sequential flow in HCI. They are relatively easy and inexpensive to set up, compared with actually building a system. They make it possible to alter a planned system's characteristics in major ways, and to study the impact of different interface features in a systematic and scientific matter (Oviatt, 2012, pp. 412–413).

The prototypical ITD environment consisted of two commercial multimodal interfaces: a 10" Android tablet (Samsung Galaxy) and a 27" Windows touch-screen computer (HP Envy Recline all-in-one), as illustrated in Figure 9 below:

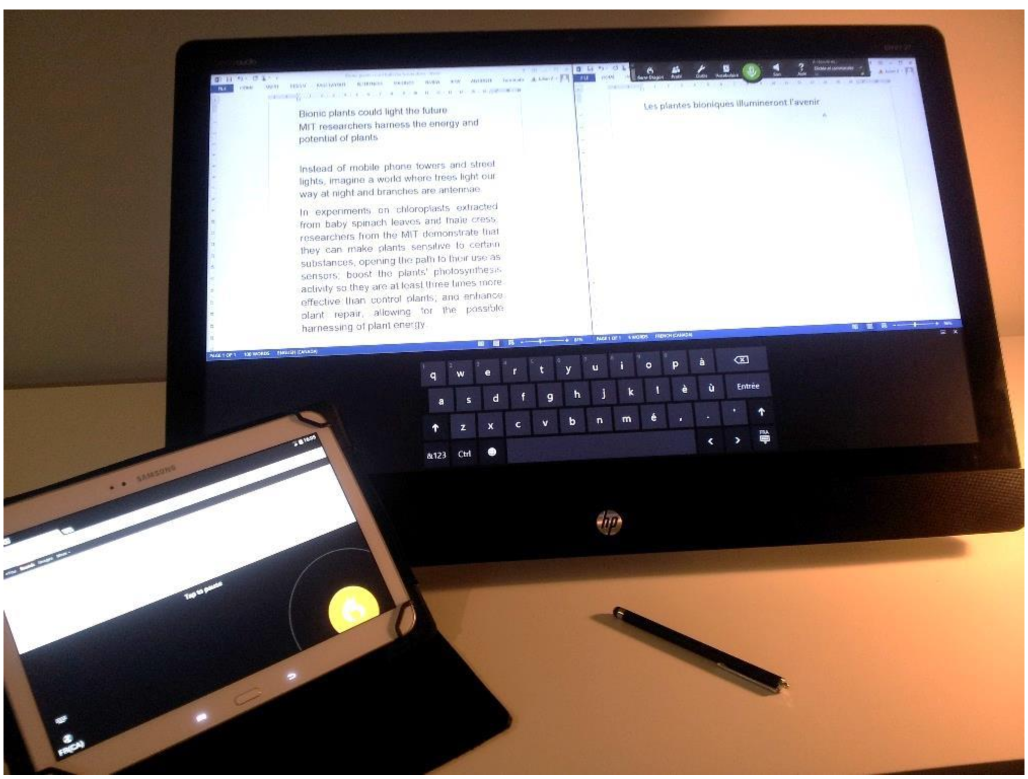


Figure 9. Physical prototype of an interactive translation dictation environment.

The touch-screen computer was equipped with a word-processing application (Microsoft Word 2013) with a CAT-specific add-on installed (a toolbar[60] developed to launch queries to informational resources directly from MS Word's interface); a state-of-the-art SA VR system (Dragon NaturallySpeaking 13 by Nuance, French version); and Windows 8's virtual keyboard (or software keyboard). Participants used this environment to read/prepare the source text and to dictate and/or type the target text and make corrections or changes to it. In addition, by means of voice shortcuts, they were able to launch queries to two informational resources directly from the source-document in Word using the above-mentioned toolbar (i.e., by selecting with the finger or the stylus a given term or character string and saying[61] "*press control eight*" to launch Termium Plus; and "*press control nine*"[62] to launch Linguee, searching the term or character string directly; a separate smaller window with the search results pops up on the screen, overlapping the Word document, as illustrated in Figure 10 below).

[60] Terminotix Toolbar for MS Word, developed and made available for free by Terminotix, Inc., a Canadian CAT tool provider. The toolbar can be downloaded from the following site: http://www.terminotix.com/index.asp?content=brand&brand=13&lang=en. Retrieved on 02-03-2016.

[61] Actually, shortcuts in the Terminotix toolbar are meant to be launched with the keyboard. Users may customize shortcuts as they wish from the toolbar's settings. The voice shortcut is possible only because the VR software has a keyboard-shortcut-by-voice feature; I combined the capabilities of both tools in the prototype.

[62] The commands were actually issued in French: "*touche contrôle huit/neuf*".

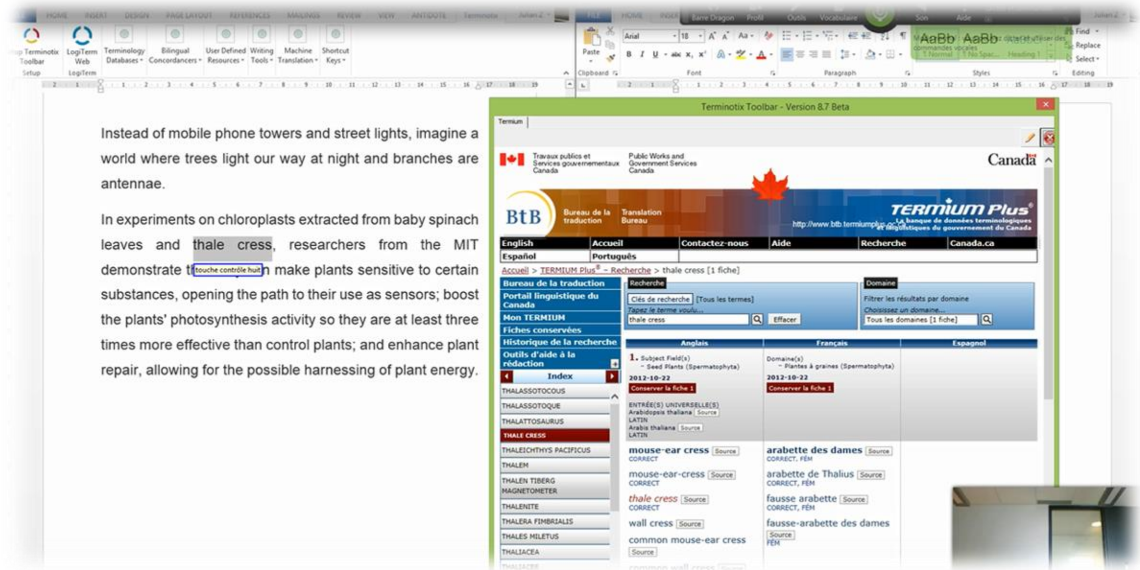


Figure 10. Screen capture illustrating the Terminotix Toolbar window with search results.

A tablet was also included in this experiment for three main reasons: (1) it provides the user with different approaches to VR technology and virtual keyboard input, as I will describe below; (2) it allows the user to reflect upon the advantages and challenges of using (smaller) mobile MIs and cloud-based applications to carry out translation tasks; and (3) it is an external platform for the user to carry out any research in any source on the web, or store any information as necessary.

The tablet was equipped with a cloud-based SI VR system integrated in a fast-typing virtual keyboard app (Swype, by Nuance) with swiping capabilities (see footnote 53 on page 87), absent in the PC's virtual keyboard. Participants in the test group used this interface to do searches online (e.g. by using Google Search, either manually in

Google Chrome or using the Voice Search app)[63] and access online tools, such as terminology databases, dictionaries and corpora, while preparing, producing and/or revising their translation. Participants could also use the Notes app to store/write down any information as they found necessary.

Figure 11 below depicts the experimental setup, and displays the hardware and software involved in the experiment. The model of each device seen in the figure is the actual model used in the experiment.

[63] I was not aware of the multilingual voice search capabilities of the Google Voice Search app, as described in chapter 4, at the time of the experiment. This being said, I did explain to the participants how to quickly switch languages in the Swype + Dragon app in case they needed to type or dictate something in a language or the other (i.e., English or French).

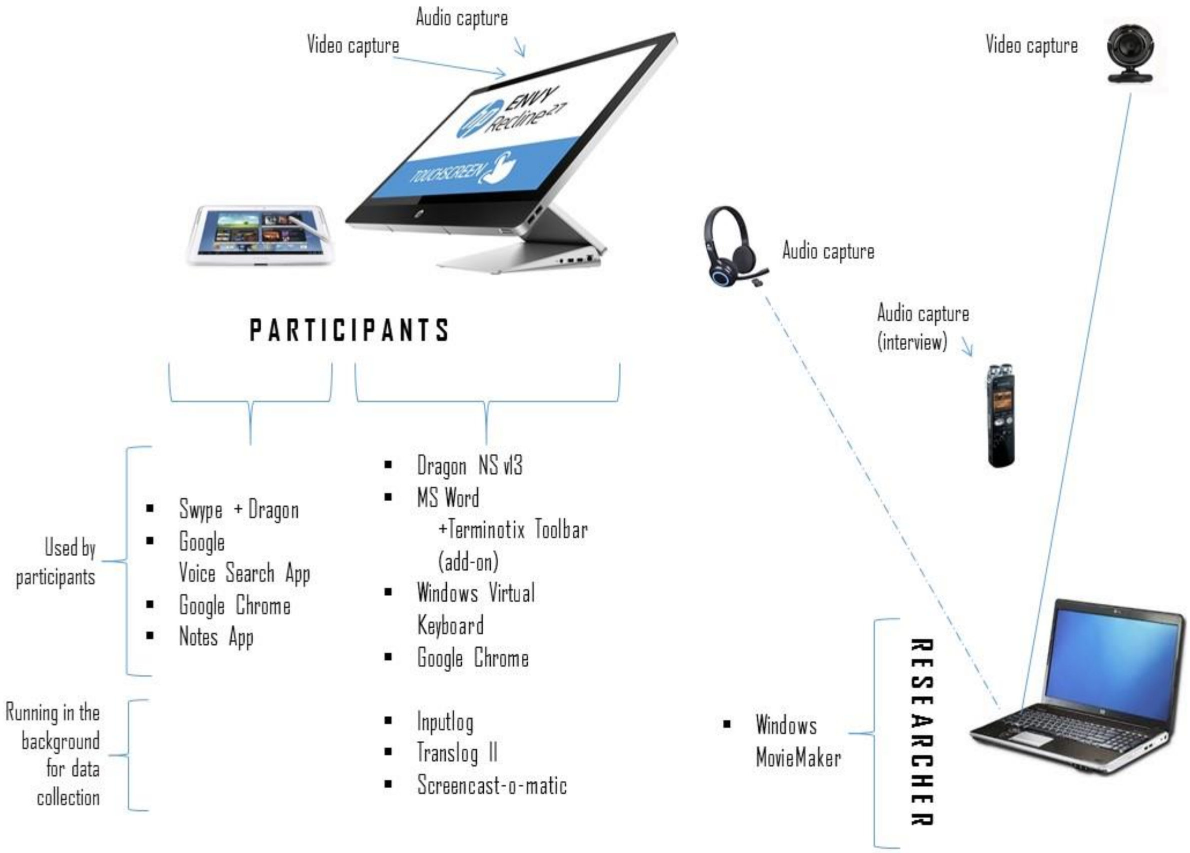


Figure 11. Experimental setup. Hardware and software involved in experiments.

Once the pre-control task was completed, participants in the test group (n=7) were given, first of all, a 15-20-minute introduction to the experimental environment. The introduction consisted of the following: first, I launched Dragon NaturallySpeaking 13 and created a user profile together with the participant (I walked the participant through the profile-creation steps). The participant had the opportunity to test the software transcribing a short simple text in French (displayed on the tablet's Notes app) by "respeaking" it into a Word document in the PC. Then, they were given the opportunity to test the touch-screen, using either their fingers or the stylus, and try to point, place and move the cursor, tap and double tap, and select, resize and move windows, etc. Then, they were given a sample translation task (a very short text in English to be translated into French), and then were given the opportunity to test the two voice commands set to launch queries to informational resources from the Terminotix Toolbar (i.e., by selecting a given term or character string and saying, in French, "*press control eight*" to launch Termium Plus; and "*press control nine*" to launch Linguee). Lastly, participants were introduced to the Google Voice Search app on the tablet, as well as the Swype+Dragon app that could be used to "type" (by swiping one's finger on the virtual keyboard) or dictate (on any text field) on this device. By the end of the introductory section, participants in the test group were instructed to perform the main experimental task: to translate a 100-word text from English into French using the prototypical ITD environment.

The second group of participants (n=7) performed the control task first. This task consisted in translating the same 100-word text using conventional input modes (i.e., a physical keyboard and a mouse). Participants were instructed to translate the text "as they would normally do" at work. It is important to note that the computer used for this task was the same all-in-one touch-screen computer used for the experimental task. However, the (wireless) keyboard and mouse that come with the computer were made available to the participants first; thus, participants were not necessarily aware at this point that the screen was touched-enabled or that there were any VR or software keyboard available. Let us also note that the only "new" feature that was added to this task (which differed from their regular workflow) was the Terminotix toolbar shortcuts. In this case, the shortcuts to Termium Plus and Linguee were not voice shortcuts but keyboard shortcuts instead (i.e., users had to actually press CTL+8 to launch Termium Plus and CTL+9 to launch Linguee). After the translation task, participants in the control group were given an introduction to the experimental environment during 15-20 minutes; the same introduction given to participants in the test group (described above).

Lastly, all 14 participants took part in a semi-structured interview at the end of the experiment for around 30-45 minutes. The whole experiment did not exceed 1.5 hours with any of the participants. Attached in appendix B are the texts used for the pre-control and main tasks and the interview guides used during the interviews with both groups.

Let us note that the interviews were designed to learn more about the participants' professional background and to provide a bigger picture of the TX with the experimental environment versus the traditional KM environment. Instead of focusing only on the tasks performed the day of the experiment, the questions were designed to provide data about the participants' perceptions of their current work environment and, in the light of the experiment, have participants reflect upon the advantages and challenges of emerging technologies and interaction modes to be integrated in the near future. I opted for semi-structured interviews instead of questionnaires in order to engage in a conversation with participants and gain further insights about the TX with the experimental and control environments. Semi-structured interviews also allow researchers to ask context-dependent questions that may concern certain participants based on their profile or on observations from their interaction experience. However, I acknowledge that subjective usability measures could have been enhanced by the use of standard pre- and post-use questionnaires, or a survey, not included in the experimental design. With data from my semi-structured interviews, it is practically impossible to calculate response rates. Not all participants were asked the same questions the same way, and conversations often went in different directions.

5.3.3. Data Extraction, Storage and Management

Having completed the data collection, the first step was to extract quantitative input data collected with Translog II for the pre-control task and with InputLog for the experimental and control tasks. The logging files provide rich interaction process data such as task times, number of characters typed with a physical or a software keyboard, number of clicks (or touches on the screen), number of switches between mouse and keyboard, time spent in specific windows, etc. Let us remind the reader that input data collection was possible with the Windows PC, but not with the Android tablet (input log software is not available for this device), which constitutes one of the main gaps of my experimental setup. However, most participants focused their attention on interacting with the PC and used the tablet only minimally.

All quantitative data was stored and managed using MS Excel spreadsheets. Graphs were generated using MS Excel as well.

The interview sessions were transcribed first of all (by respeaking them using a VR program in French). Given that interview answers are spontaneous statements formulated by the interviewees, it was necessary to restructure the answers in a coherent matter, not only to improve the accuracy of transcription with VR, but also to improve readability. When answering an interview question, a person may not organize his or her thoughts coherently before formulating an answer. Therefore, I took special care to reformulate the answers in a syntactically and grammatically correct language, while taking special care not to leave out any important information. As an example, an exact utterance by one participant was as follows: "*Quand… quand j'avais essayé tantôt, c'est… au début j'avais d'la misère puis après ça... ça allait mieux, là, quand j'ai compris qu'il fallait pas lever le… puis qu'il fallait arrêter entre chaque mot, là; au début j'étais pas capable de l'faire, mais à la fin ça… ça allait assez bien, là.*" After transcription, reformulation and translation, the same utterance became this: "I ended up doing pretty well with the Swype keyboard once I understood that I did not need to lift up my finger, and that I needed to stop after each word." Lastly, it was necessary to remove any information that may reveal the participants' identity or that of their organization.

When the transcription step was completed, the data was imported into NVivo10[64], a computer-assisted qualitative data analysis program. NVivo allows researchers to classify, sort and arrange rich and complex qualitative data, and examine relationships in the data; it is an ideal tool to organize and analyze data such as interview answers and surveys, for instance. Interview data were classified into various categories according to the different aspects of usability measured (e.g. perception of task difficulty, perception of the outcome, subjectively experienced

[64] NVivo qualitative data analysis Software; QSR International Pty Ltd. Version 10, 2012. A newer version of the software is now available as of December 2015. See http://www.qsrinternational.com/ for more information.

time) and other context-dependent questions that emerged during the interviews (e.g. description of current working environment, levels of job satisfaction, tool-specific and external difficulties for ITD integration).

To obtain further objective measures, the 14 translations were submitted anonymously to three expert evaluators (translation professors at the University of Ottawa). Although evaluating a translation could be considered a subjective task, the average of the three grades provides us with a more objective measure of the quality of the product of the interaction. Evaluators were instructed to mark the translations using the criteria they use for class assignments and evaluations. No details on the participants' profile or the experimental setup were given to the evaluators.

5.3.4. Challenges of The Study's Methodology and Experimental Design

To the best of my knowledge, an empirical study involving off-the-shelf multimodal interfaces in translation is unprecedented. No other study in TS has investigated the use of VTS input and comparing it to KM input. In addition, only a limited number of experiments in TS have used various data-collection methods such as video recording, screen recording and interviewing. Thus, there is not much literature in TS upon which to build a solid methodology for this experiment. Indeed, I was using an array of devices and software that had not been combined before in one translation experiment. Despite several pilot projects, preliminary tests of the experimental environment and simulations, many unexpected issues (e.g. system breakdowns, loss of internet connectivity) were encountered during the experiment sessions, as well as during the analysis phase.

Another major drawback to be noted in my experimental design is limited input logging data. As it was discussed in the methodology overview above, participants in the test group were interacting both with a touch-screen computer and with a tablet, but screen recording and input logging was running only for the touch-screen computer. To the best of my knowledge, Android devices, or at least the tablet I was using, do not support input logging and screen recording software. In addition, the input logging software I used on the touch-screen computer was not specifically developed for logging voice-and-touch interactions. An attempt was made to combine logging data from the VR application and the input logging application but the process failed repeatedly; these data was therefore discarded from the analyses.

Regarding evaluations, it is considered as a methodological limitation the fact that objective measures of quality were obtained from academics (university professors), and not professional revisers (whose daily job is to assess and improve the quality of translations performed by their peers). Professors are likely to approach the evaluation in a different way than revisers would, surely with objectives that are *a priori* pedagogical. All three evaluators in my study normally use the same evaluation grid, even though, as stated above, their perception of the quality of a translation may be subjective. This being said, generally speaking, no differences were noted between their evaluations.[65]

[65] A detailed analysis and discussion of the differences in evaluations is beyond the scope of this study, but would be essential and ideal in studies that look more closely at measures of the outcome of the interaction such as completeness and precision, not included in this thesis.

It is also considered as a major limitation of this study the fact that all participants were first-time users of the experimental environment. Only a short introduction was provided. Ideally, a longitudinal study (including training sessions) would have been carried out to provide a better answer to the research question, but such a study was not possible due to the many limitations specific to a doctoral study in the humanities, as described in section 1.6.

Another methodological limitation was the length and the nature of the text used for the translation task. The data are based only on the interaction with the experimental and control environments to translate a 100-word text adapted from a news article, which is far from representative of real-life assignments participants perform at work on a daily basis. In sum, not much can be extrapolated, hypothesized and concluded from such a short text.

Lastly, I opted for semi-structured interviews instead of questionnaires in order to engage in a conversation with participants and gain further insights about the TX with the experimental and control environments. I acknowledge that subjective usability measures could have benefited from the use of standard pre- and post-use questionnaires, or a survey, not included in the experimental design.

It is also worth noting at this point that, while I acknowledge the importance of TMSs in current translation workflows, this technology was excluded from the experimental portion of my study. First of all, an experiment that includes this technology and emerging technologies such as voice and touch input is rather challenging and goes beyond the scope of this thesis. In addition, in order to create a basis for comparison, I am exploring ITD to produce from-scratch translations only, since TMSs are not appropriate in some instances and contexts. However, I do not rule out completely the inclusion of TMSs in an ITD environment, as they may indeed augment the workflow, and improve the efficiency of the ITD system in the background. Such an exploration will remain for future work. The use of TMS was nonetheless part of the discussions with participants during the interviews (see Chapter 6 for more details). I also refer the reader to a recent ethnographic study by Leblanc (2013) to learn more about professional translators' perceptions and opinions on the use of TMSs.

The sample, as well as the tools and methods used, sufficed for the purposes of this research, as the main goal is to build a foundation for future TCI and TII research. The aim was also to collect preliminary user data to design a functional prototype of an ITD environment in the near future. A longer-term study including a larger sample and a bolder methodological design would be the next logical step.

The following chapter presents a detailed analysis and a discussion of the quantitative data collected in the pre-control task, on the one hand, and in the experimental and control tasks, on the other hand; as well as the qualitative data from the interviews.

Chapter 6 — Translator Experience Data Analysis and Discussion

"We think that if we have made a clever and thoughtful argument, based on data and smart analysis, then people will change their minds. This isn't true. If you want to change people's behaviour, you need to touch their hearts, not just win the argument." — Eric Schmidt

This chapter provides an analysis and a discussion of the TX data collected during the experiments. It first provides a description of the participants' profile. Then, it presents the quantitative data from the pre-control, control and experimental tasks, and finally the qualitative data collected during the interview.

In order to maximize anonymity, the data are presented in an aggregate format; no links are made between the participants' metadata and their experimental results, and any information that may reveal the participants' identity or that of their organization has been left out from the analysis presented in this thesis. Given this methodological limitation, no generalizations can be formulated from the findings of the present study.

6.1. Participants' Profile

The sample (n=14) consisted of professional translators working in a Canadian translation organization. They had at least 3 years up to 33 years of experience in translation by the time of the experiment.

Participants were divided into two groups: a test group (n=7) and a control group. Variables such as age and years of experience were counterbalanced to ensure, as much as possible, an equal distribution of participants into the two groups.[66] Table 2 below shows the number of participants per age group:

26-30	4
31-35	3
36-40	4
41-45	0
46+	3

Table 2. Number of participants per age group.

The sample included 3 men and 11 women. However, in this thesis, masculine pseudonyms are given to all participants in order to ensure uniformity (e.g. in the use of personal and possesive pronouns while presenting results). Another reason for this choice is because there were fewer men than women; therefore, by giving masculine pseudonyms, I believe to be maximizing their anonymity. Please note that this choice was completely personal and arbitrary, and was made for no other reason than those just mentioned.

[66] This distribution was carried out with the help of an expert statistician prior to performing the experiments. Using the R statistics software environment, as well as metadata collected via e-mail from participants a few days before the experiment, it was possible to place participants in the two subgroups. For more information on the R software, please see: https://www.r-project.org/. Retrieved on 16-06-2016.

Table 3 below shows the participants' distribution into the two groups, and the participants' codes and pseudonyms to be used in this thesis.[67] Both codes and pseudonyms will be used indistinctively in this discussion.

C/T	Pseudonym
C1	Carlos
C2	Pablo
C3	Pedro
C4	Sebastián
C5	César
C6	Iván
C7	David
T1	José
T2	Andrés
T3	Mauricio
T4	Jaime
T5	Alejandro
T6	Antonio
T7	Lisandro

Table 3. Participants distribution and pseudonyms. C= Control group; T= Test group.

Although all participants reported to have at least 3 years of professional experience in translation, those who currently hold management positions (see Participant Recruitment and Sampling, section 1.6.2) had not worked doing translation per se full time for at least one year and up to 13 years (average = 4.8 years) at the time of the experiment. This variable was also taken into account when dividing participants into the two groups. Together with the huge gap in the number of years of experience

[67] These pseudonyms are different to those used during the experiments. Each participant was given a pseudonym according to their gender during the experimental session. No real name was used at any time. Now, not even participants themselves should be able to recognize their own experimental results or those of their peers based on the new pseudonyms.

(3-33 years), this gap makes it difficult to formulate any generalizations from the findings of the experiment.

All participants reported to have French as their mother tongue. Eight participants reported that they acquired the English language in early childhood, and the other six as teenagers. Likewise, six participants reported that they are able to translate from English into French and from French into English indistinctively. In addition, four out of fourteen participants reported to be left-handed; the other ten reported to be right-handed.

Eight participants reported that they own a tablet and ten reported that they own a smartphone; six participants own both. In addition, out of 14 participants, only one said he owns a touch-screen computer, but that he never touches the screen. Let us also mention that two translators in our sample own neither device.

Lastly, half of participants had not used VR before, the other half had used it minimally either for work (doing online TDVR; see Chapter 3), or for issuing commands to their smartphones or in-vehicle navigation system.

6.2. Quantitative Data Analysis

Quantitative data analysis constitutes the first step towards confirming or falsifying my hypothesis, that is, whether or not the prototypical ITD environment consistently provides translators with a better TX than the KM environment. Various measures will be presented and discussed here: total task times, time-in-mode, deviation from optimal solution, usage patterns, input rates, and expert assessment of the outcome of the interaction.

6.2.1. Pre-control Task

The pre-control task allowed me to collect data on the participants' typing speed, transcription accuracy, reading times and other input data such as keyboard and mouse activity. Some objective interaction process measures (see Table 1 on page 91) are used in the analysis.

6.2.1.1. Total Task Times and Time-in-Mode

Using the Translog II log files, it was possible to extract and compare the total time spent by the 14 participants in the document versus the actual time spent typing (i.e., in active writing mode). Figure 12 below displays the typing times per participant, and average. In this figure, the following can be observed: (1) Participants may have different typing behaviour, and (2) participants' typing speed may differ considerably between them. First, it may be seen that for some participants the total time spent in the document is only slightly higher than the total time spent actually typing. This may be partly explained by the fact that some participants are touch typists and need not switch their attention from the screen to the keyboard, or from the original text to the text field where they are typing and to the keyboard again. However, this time difference may be considerable in some instances, as it was the case of participants Pablo (C2), César (C5) and Alejandro (T5). Switching their attention from the keyboard to the screen may have a significant impact on their productivity, and that, carrying out a simple transcription task. Likewise, Figure 12 also shows different typing speeds among participants. Participant Pedro (C3), for instance, who is the fastest typist in the sample, could type the text almost 2½ times faster than César (C5), the slowest typist, although they were all transcribing the same text and were doing so in their mother tongue (but still considering that they were not using the keyboard they are used to typing with at work or at home). The keyboard used for this task was a Canadian multilingual standard keyboard set up for the French (Canada) language in the Windows language bar. Before carrying out the pre-control task, participants were given a minute or two to become acquainted with the keyboard (location of accented and special characters, etc.) until they said they were ready for the task. All participants reported that this keyboard is different to the one they are used to working with at home or at their office, and acknowledged that this might have had an impact on their performance. The variability of computer keyboards is to be explored in future research in this vein.

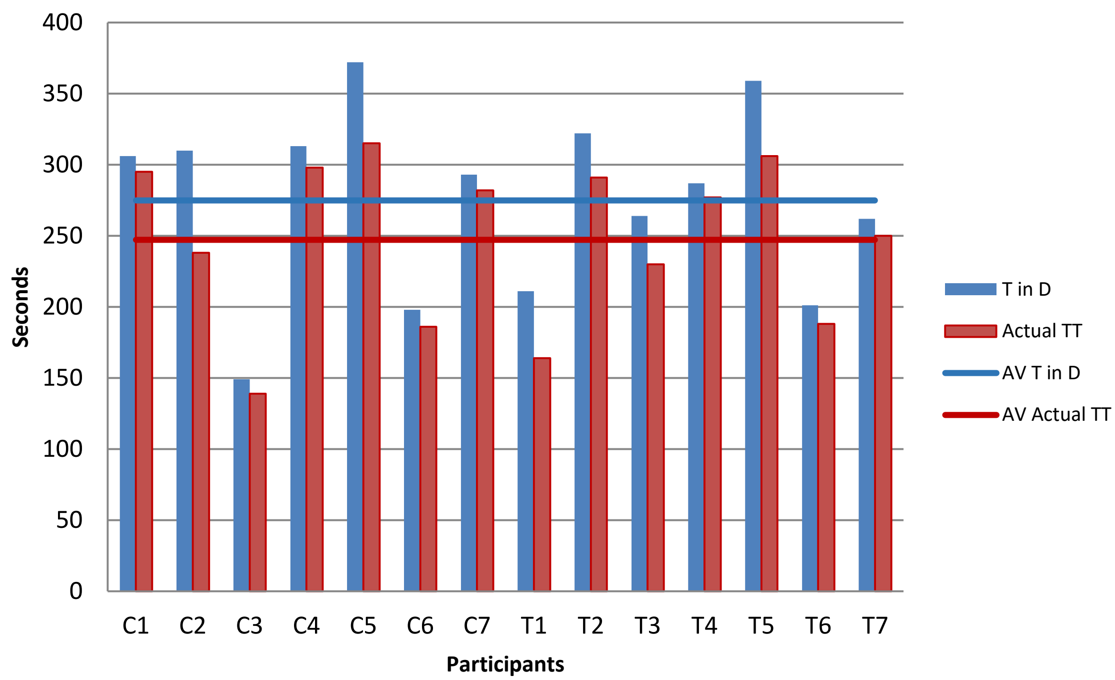


Figure 12. Typing subtask times (in seconds) per participants and average (AV). T in D = total time spent in the document; Actual TT: total time spent in active writing mode.

In sum, a person's ability to type with a physical computer keyboard (or the keyboard itself, if different from the keyboard the person is used to working with) may have an impact on their performance while carrying out simple linguistic tasks such as copying a text, and this may be a key factor also when comparing their performance with that of their peers.

Now, how much faster can a person read aloud a text than he or she can type it? Figure 13 below displays participants' reading aloud times in relation to their own typing speed. From this figure, it can be concluded that the 14 participants take approximately the same amount of time to read aloud a text in their mother tongue, whereas typing it may take from nearly three to over six times longer (over four times on average).

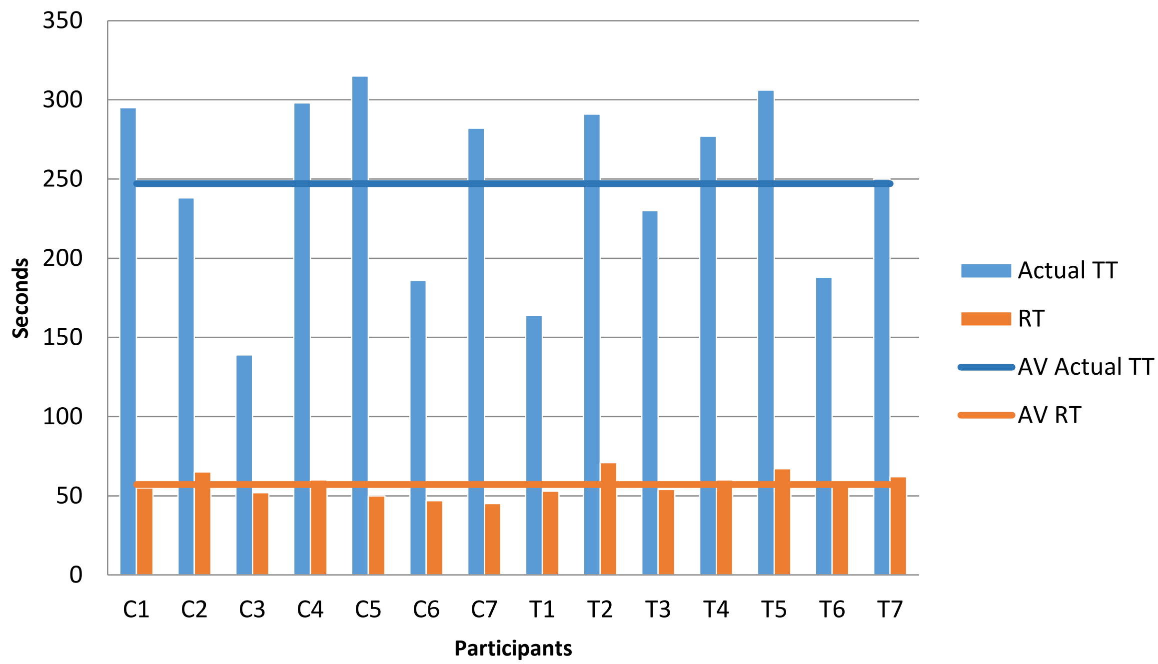


Figure 13. Time in active writing mode (Actual TT) vs. reading time (RT) (in seconds) per participant and average (AV).

In short, even the fastest typists from our sample, i.e. Pedro (C3) and José (T1), need approximately three times more time to type a text than they need to read it out loud at a "normal, natural speed" (i.e., approx. two and a half minutes to type a text they can "say" in about 50 seconds). Again, a person's typing skills (and level of comfort with a given keyboard) may have a significant impact on their performance when carrying out linguistic tasks such as transcribing a text, and this may be a key factor also when comparing their performance to that of their peers. However, the participants' ability to read aloud/dictate/orally produce a text remains comparable: they take approximately the same amount of time.

6.2.1.2. Deviation from Optimal Solution

Even though this task consisted only of transcribing (copying) a short text in a parallel text field in the same window and using an identical font (the best conditions for such a task), the final product may not be error-free. The deviation from an optimal solution can be measured for the transcription sub-task.

Figure 14 below shows the transcription accuracy rates across participants for the transcription sub-task. It can be seen in fact that only two out of 14 participants achieved 100% accuracy, meaning that the final product matched exactly the original text. For the rest of participants, transcription accuracy ranged from 97.7% to 99.9% – almost perfect, but not quite.

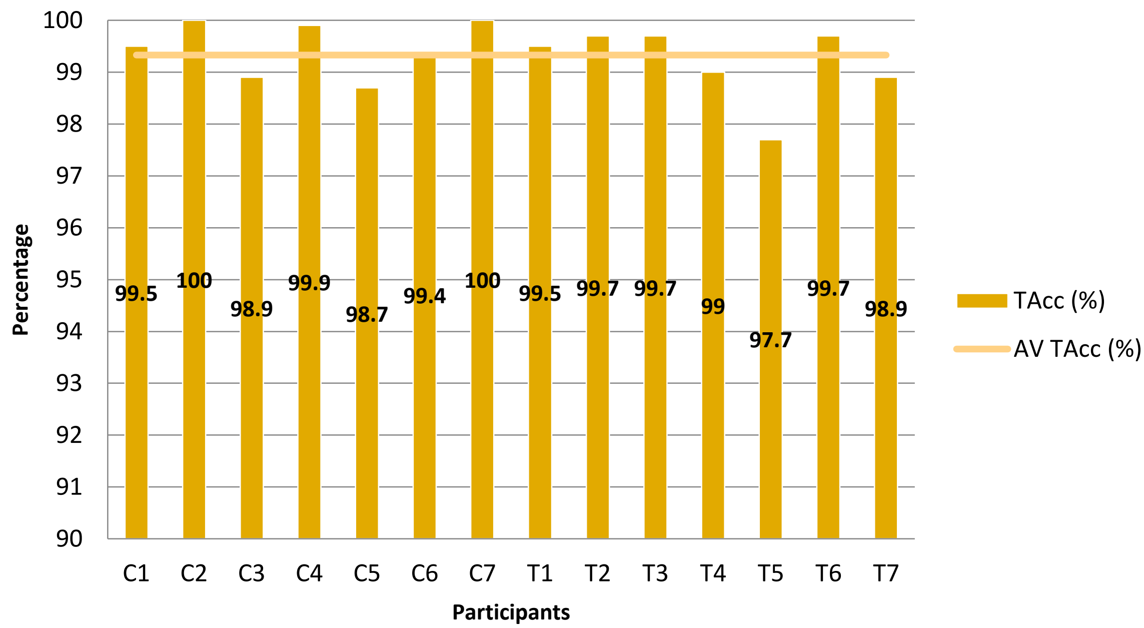


Figure 14. Transcription accuracy (TAcc) rates per participant and average (AV).

Let us note at this point that, in this research, the length of texts is measured and analyzed in terms of the number of characters including spaces, instead of the number of words. The number of characters including spaces is a more precise measure than the widely used text-length measures by the number of words. In this study, I am interested in investigating the technical and physical effort needed to transcribe texts, and such investigation can be carried out more accurately at the character level.

To calculate the transcription accuracy shown in Figure 14 above, a simple online edit distance calculator (also known as Levenshtein edit distance (Navarro, 2001) calculator) was used.[68] Such a tool calculates the "cheapest" way to transform one string into another. The result obtained indicates the "total cost" or, in other words, the minimum total number of keystrokes that would be needed to edit a given text (in the case of this experiment, the final product of the transcription sub-task) to match another text (in this case, the original text). That number of keystrokes is compared to the total number of characters including spaces of the original text. This gives us the percentage of the transcription accuracy. This method is also employed to calculate, for instance, the word accuracy rates of VR systems (see Zapata & Kirkedal (2015)).

In conclusion, although this task involved retyping a very short text and without any time constraints, the output of the interaction was not always error-free.

6.2.1.3. Usage Patterns

Translog II's log files provide usage pattern measures that reveal other skills and behaviour while performing a typing task, in addition to the transcription accuracy rates shown in Figure 14 above. Usage patterns measures obtained include total user

[68] The tool, developed by Peter Kleiweg, is available for free at: http://odur.let.rug.nl/kleiweg/lev/.

events, text production events, text elimination events (i.e., deleting), and miscellaneous events (which may include, in this case, mouse clicks or scrolling with the mouse wheel). Table 4 below reveals for instance that none of the participants' typing is 100% accurate (i.e., they needed more keystrokes than the total number of characters including spaces of the text) and that they all have different typing skills and behaviour: on average, to type an 862-character text, participants needed 911 keystrokes in active writing mode (text production) including spaces, 51 deletions (text elimination) and 19 clicks (miscellaneous events), for a total of 982 events. Note that participant Pedro (C3) is the most skilled participant with the keyboard. Not only is he the fastest typist (see Figure 12 above), but also needed the least "technical effort" (i.e., the total amount of keyboard (and mouse) events) to transcribe an 862-character text: 875 keystrokes in active writing mode, 15 deletions, and 4 mouse clicks, to arrive nonetheless at a transcription that is 98.9% accurate.

	TUEv	TxT P	Txt E	MEv
C1	926	889	31	6
C2	1010	905	45	60
C3	894	875	15	4
C4	1113	979	119	15
C5	1025	935	80	10
C6	1006	920	59	27
C7	1070	953	93	24
T1	946	900	40	6
T2	1017	920	60	37
T3	910	877	17	16
T4	986	923	59	4
T5	986	898	38	50
T6	962	909	49	4
T7	905	879	20	6
AV	**982.57**	**911.57**	**51.79**	**19.21**

Table 4. Event measures. TUEv = total user events; Txt P = text production; Txt E = text elimination; MEv= miscellaneous events; AV= average.

6.2.1.4. Input Rates

Translog II also allows a researcher to measure input rates. Figure 15 below shows the text production speed versus the total user event production speed (that is, including text elimination and miscellaneous events) per minute. Input rate data may allow a researcher, for instance, to calculate the minimum amount of time a participant would take to transcribe a text (provided that the typing is 100% accurate). Taking again the example of participant Pedro (C3), who is the most skilled typist, based on his data, if his typing were 100% accurate, it would still take him over 2.5 minutes to transcribe the same text that takes him 53 seconds to read aloud (see

Figure 13 on page 121); in short, with perfect typing skills, typing would still take at least 2.5 times longer than speaking.

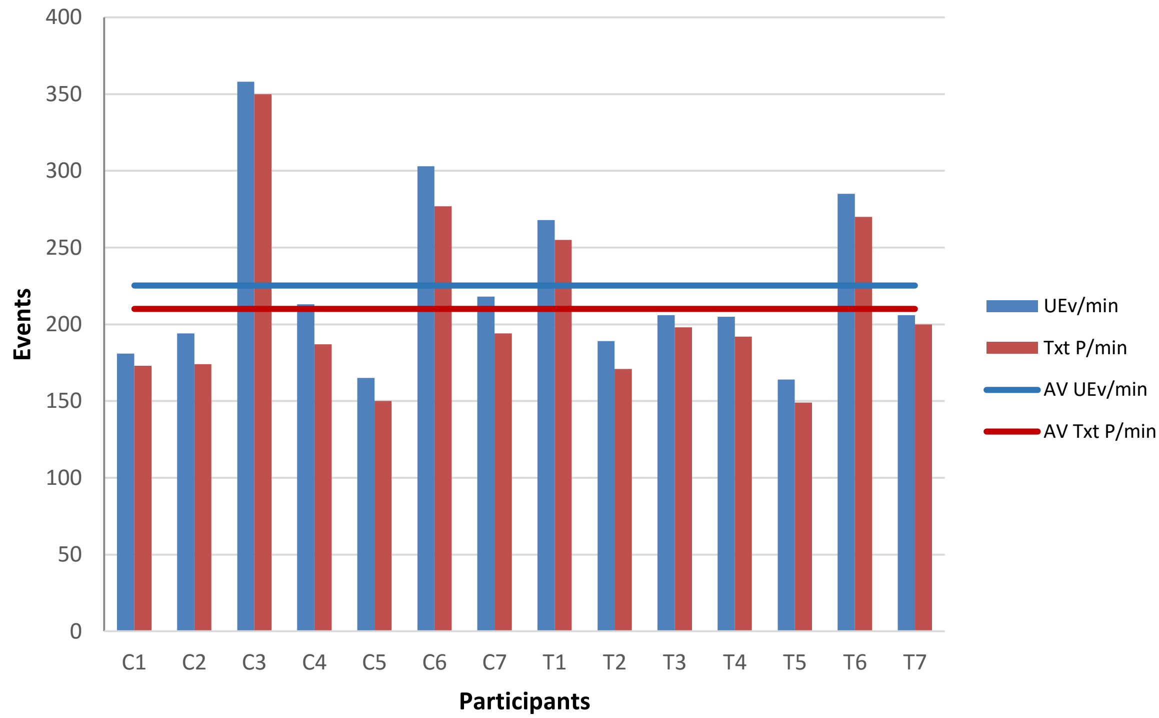


Figure 15. Total user events per minute (UEv/min) vs. Text production events per minute (Txt P/min) per participant, and average (AV).

As we can see, input logging data can be rich and relevant for analyzing text production behaviour and skills. Data provided by Translog II was adequate for this pre-control experiment, which yielded sufficient results for investigating the participants' skills and behaviour with traditional input devices, i.e., with the keyboard and mouse, based on a simple copying task in their mother tongue.

Translog II was used for the pre-control task, since it involved a simple writing task. For the main experiment, however, a more powerful application was needed to collect richer interaction data. The Inputlog program provides a more complete dataset and allows HCI, HII and writing process researchers (and in the case of translation, TCI, TII and translation process researchers) to formulate better hypotheses about computer- and information-interaction behaviour. Inputlog offers system-wide logging capabilities, allowing researchers to track any application running in the operating system (Teixeira, 2014), from text-processing applications to informational resources. Let us now look into the experimental tasks results in detail.

6.2.2. Main Experiment

In this section, the quantitative data collected from expert evaluators and with Inputlog software during the main experiment is presented and discussed. Some objective outcomes and interaction process measures (see Table 1 on page 91) are used in the analysis.

6.2.2.1. Expert Assessment of Outcome

First of all, prior to extracting and analyzing input data, the first phase of the quantitative data analysis consisted of obtaining and analyzing objective measures of quality. To obtain these measures, the 14 copies of the translations were submitted anonymously to three expert evaluators, who are translation professors at the University of Ottawa. Tables 7 and 8 below display the alphabetic grade given by each evaluator to the participants' translations in the control group and the test group respectively, and the average numeric grade as per the University of Ottawa grade conversion table:[69]

nPseudo	EV1	EV2	EV3	AV
Carlos	C+	B	C	**5.0**
Pablo	A-	B+	A-	**7.6**
Pedro	C+	*D+*	B+	**5.0**
Sebastián	C+	A-	B	**6.3**
César	D+	D	C	**3.0**
Iván	B	A	C	**6.3**
David	A+	A	B+	**8.6**
AV				**6.0**

Table 5. Expert evaluations in alphabetic grades per participant in the control group, and average numeric grade.

nPseudo	EV1	EV2	EV3	AV
José	B	C	B	**5.3**
Andrés	B	B+	B	**6.3**
Mauricio	B	B+	B	**6.3**
Jaime	C+	D	C	**3.6**
Alejandro	B	B+	A-	**7.0**
Antonio	B	*C+*	B+	**6.0**
Lisandro	B	*C+*	B+	**6.0**
AV				**5.8**

Table 6. Expert evaluations in alphabetic grades per participant in the test group, and average numeric grade.

[69] The alphanumeric table is available for consultation at the following site: http://www.uottawa.ca/administration-and-governance/academic-regulation-10-grading-system. While analyzing the data, I realized that one of the evaluators attributed, in 3 out of 14 of cases, grades that do not appear in this table (B- and C-), although they confirmed that they were using the official grid. I considered this to be an honest mistake by the evaluator. In my analysis, I established that B- was equivalent to C+ (5/10), and that C- was equivalent to D+ (3/10). These 3 cases have been indicated in italics in the tables.

It was found that the average grade for the two groups was comparable (out of 10 points, 6.0 for the control group and 5.8 for the test group); thus, based on this sample, the experimental setup has a minor impact on objective quality when comparing with the quality of the outcome with the control environment. This being said, by taking a closer look at the grades, it was noted that the quality of the control group varies considerably (from the lowest grade, i.e. 3.0 [participant César (C5)] to the highest, i.e., 8.6 [participant David (C7)]), whereas the quality for the test group seems more stable, with the exception of participant Jaime (T4), whose grade was 3.6 out of 10 points.

As stated in the previous chapter, all three evaluators used the same evaluation grid. Except for a couple of cases, no considerable differences were noted between the evaluations. A detailed analysis and discussion of the differences in evaluations is beyond the scope of this study, but would be essential and ideal in studies that look more closely at measures of the outcome of the interaction such as error rates, completeness and precision, not included in the analysis presented in this thesis.

6.2.2.2. Task Times

As far as task times are concerned, total translation times are presented first, that is, how long it took participants to do the translation from start to finish. Although Inputlog logging files provide total task-time data, it was necessary nonetheless, with the help of the screen recordings and a stopwatch, to subtract the time at the beginning and at the end of the task that was not part of the translation task per se (e.g. preparing the environment for the experiment at the beginning, or saving the file and stopping the recording at the end. It was a matter of a few seconds, but I tried to provide data as accurately as possible).

Now, how much time do the participants take to carry out the whole translation task from the moment they start reading and preparing the source text until they finish revising the target text (and say they are ready)? Figures 16 and 17, respectively, display the total translation task times for the control and test groups, and averages. From these two figures, no particular trend seems to be noticeable as far as total translation times are concerned. First of all, it was demonstrated already with the pre-control task that participants may have varying levels of skills with the various input devices, which may be affecting their performance in terms of efficiency, and thus the total translation task times.

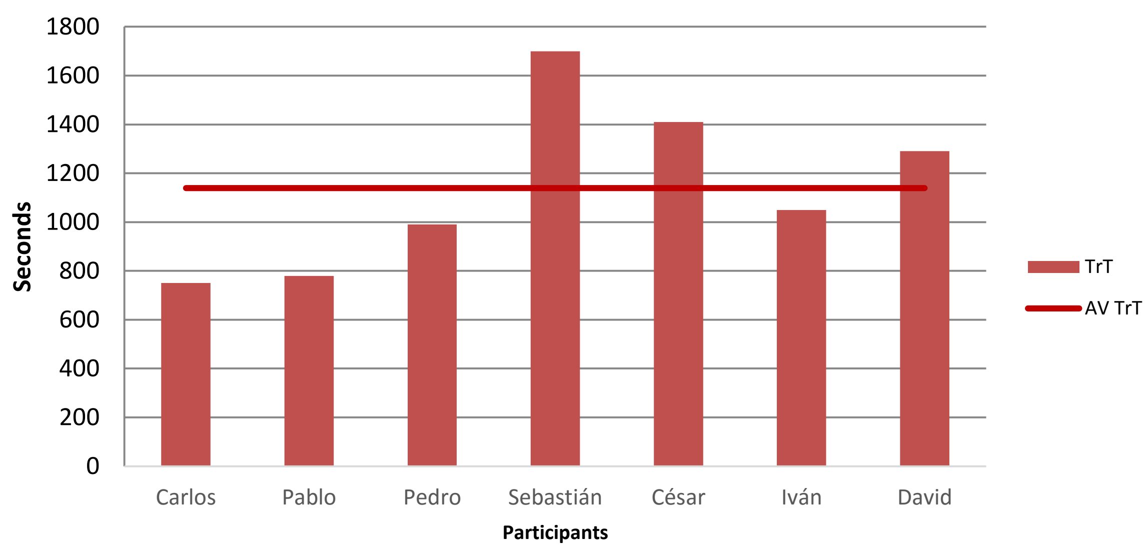


Figure 16. Translation times (TrT) (in seconds) per participant in the control group and average (AV).

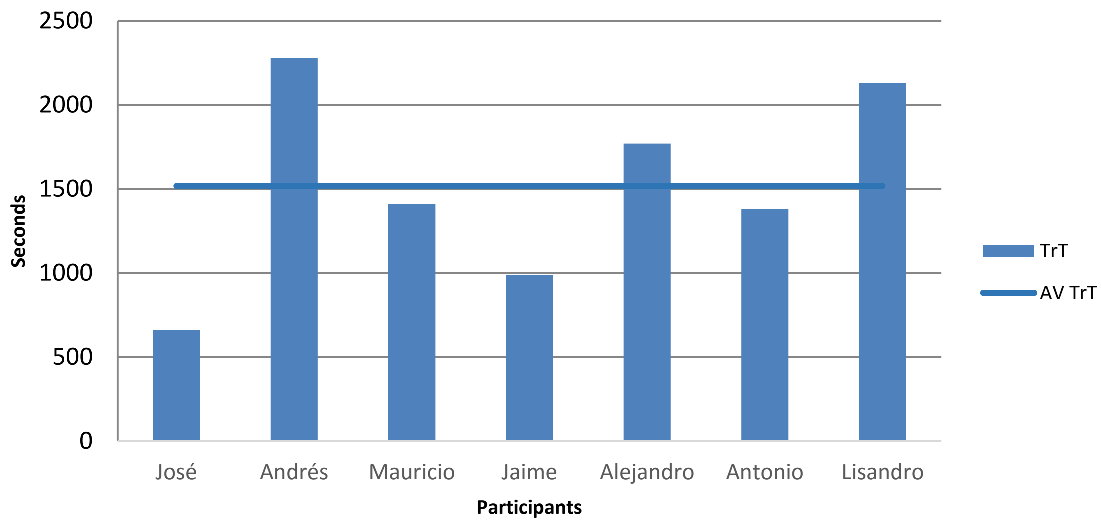


Figure 17. Translation times (TrT) (in seconds) per participant in the test group and average (AV).

It can be noted here that, contrary to what could be expected from a multimodal environment, participants in the experimental group achieved lower overall efficiency. Indeed, average total translation times were noted to be higher for the experimental group (1567 s.) than they were for the control group (1138 s.). However, based on these results, task time measures (and averages) alone are not sufficient indicators of efficiency of a translator or of a given environment, in this case, the KM environment versus the multimodal environment. In other words, it is not because participants in the test group took more time (on average) to translate the text that the prototypical ITD environment is less efficient than the KM environment. In fact, a closer look at the data reveals that both the lowest and the highest total translation times were in the experimental group (José [T1] with 660 s. and Andrés [T2] with 2280 s.). As it will be demonstrated later with the analysis of other objective measures, and subjective measures, many factors of different nature may have an impact on the participants' total translation task times and therefore on the TX and the quality-in-use of a given environment.

6.2.2.3. Time in Mode

The time spent carrying out specific tasks, with specific devices, or in specific interaction modes, is also important in the analyses.

Earlier, figures 16 and 17 displayed the total translation task times for the control and test groups, respectively. Now, how much of that time was spent *actually* typing (i.e., in active writing mode)? Inputlog provides us with an answer to that question. Figure 18 below displays the time in active writing mode with respect to the total translation time per participant in the control group, and average. From this figure, it can be concluded that when producing a translation in a KM environment, between 30 and 55% (42% on average) of the time is spent in active writing mode; the rest of the time may be spent doing peripheral tasks such as reading the source text and searching information.[70] In addition, time in active writing mode across participants remains relatively comparable (with the exception of César (C5), whose time more than doubled that of Carlos (C1), with the lowest time in active writing mode), even though their total translation times vary considerably.

[70] A detailed analysis of video and screen recordings could have allowed me to time each one of those peripheral tasks. Such an analysis is nonetheless beyond the scope of this thesis, but remains for future work.

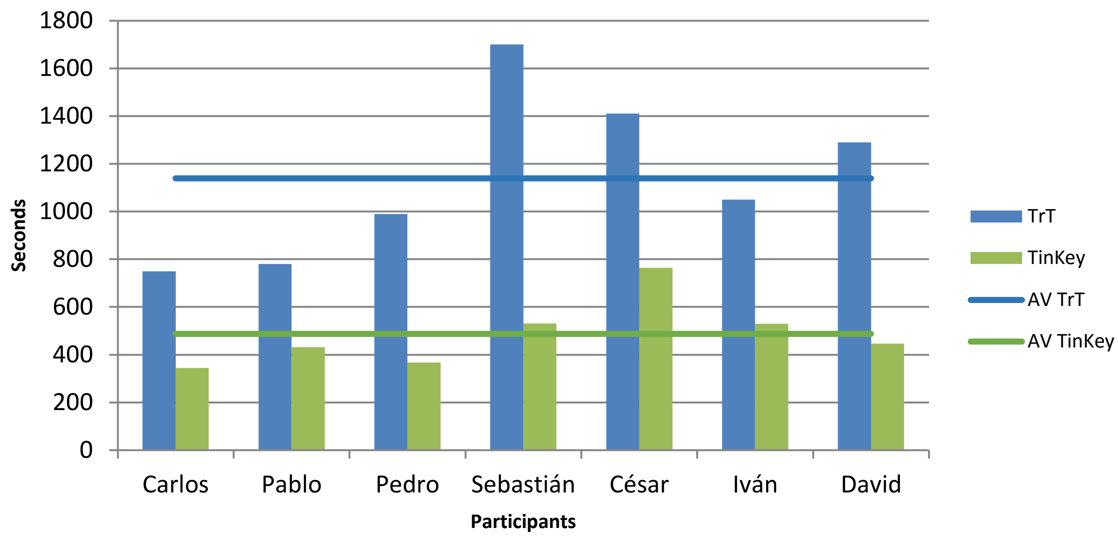


Figure 18. Time in active writing mode (TinKey) (in seconds) with respect to total translation time (TrT) per participant in the control group and average (AV).

Figure 19 below displays the time in active writing mode (i.e., typing on the software keyboard) with respect to the total translation time per participant in the test group, and average. A different trend can be observed in this figure. First of all, it is not surprising that the time in active writing mode is lower, on average, than in the control group, since participants are expected to produce most of their texts using VR, not the software keyboard. However, because this is a multimodal environment, participants had the possibility to manually intervene in the text using the keyboard as much as necessary, in combination with voice input. In other words, they were free to combine voice and software keyboard input as much as they wanted.

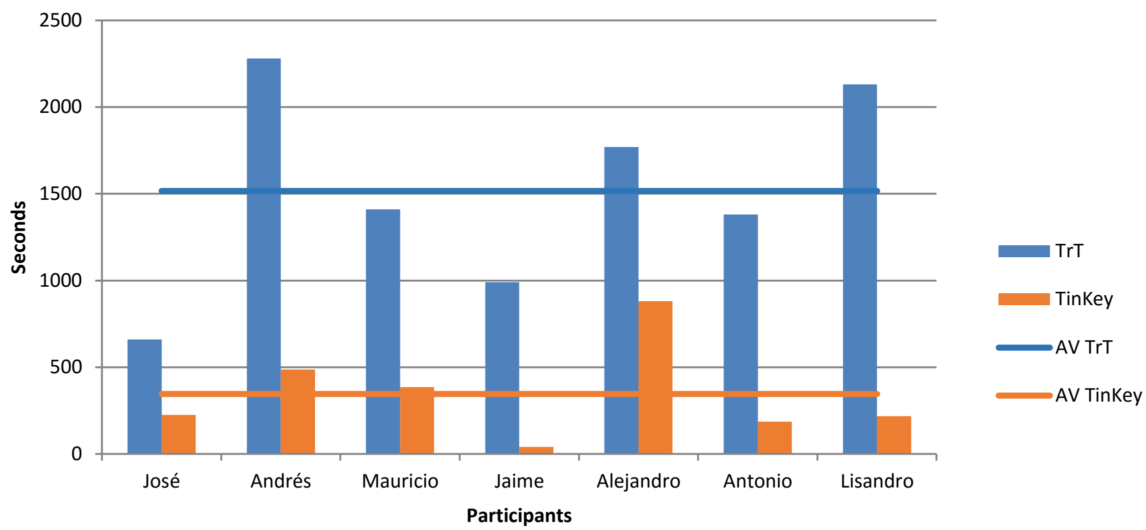


Figure 19. Time in active writing mode (TinKey) (in seconds) with respect to total translation time (TrT) per participant in the test group and average (AV).

Let us take the example of participant Jaime (T4), who spent only 4% of his total translation time using the virtual keyboard. The total translation time was below the average (and that, of both groups); however, his grade was the second lowest among all participants and the lowest among participants in the test group. Alejandro (T5), on his side, spent the most time using the keyboard (881 s.), both groups combined. This is almost 50% of his own total translation times; the other half of the time was spent reading the source text, dictating the first draft of the translation, and searching information. His total translation times (1770 s.) were only slightly higher than the average for the test group (1517 s.), but his grade was the highest in this group (7.0).

Having investigated times in active writing mode, let us now explore times with other input modes. Indeed, beyond the physical keyboard, the virtual keyboard and the VR software, other input modes are a central matter of investigation in this research, namely the mouse, on the one hand, and touch and stylus input on the touch-screen, on the other hand. Inputlog can also measure these interactions.[71]

In particular, instead of just the usage patterns, I was interested, for the purposes of this thesis, in measuring the total time spent using the mouse, or the fingers and/or stylus on the touch-screen.[72]

Figure 20 below shows the total time spent using the mouse per participant in the control group, and average. Different mouse interaction behaviour can be observed in this figure. For instance, Sebastián (C4)'s time with the mouse is more than three times that of Pablo (C2). In addition, when adding up the time spent with the mouse and time in active writing mode with the keyboard, it can be noticed that the time spent using these two input devices is considerable (again, not taking into account the number of mouse events such as clicks, or the distance "walked" by the mouse pointer on the screen, which can also be measured with input logging software). It is interesting to note, for instance, that Sebastián (C4), whose total translation task time was the highest of the control group (1700 s.), also spent the most time with the mouse. Out of 1700 seconds, 531 were spent in active writing mode, and 927 were spent either moving the mouse around, clicking, or scrolling up and down with the mouse wheel.

---

[71] Note that Inputlog was designed primarily to collect input data from KM interactions on a Windows-based computer. In other words, the program does not know the difference between clicking with the mouse and tapping on a touchscreen with a finger or a stylus. Only the researcher can tell the difference knowing in which conditions each dataset was collected.

[72] I do investigate other types of keyboard usage patterns in the subsection below. I deliberately left mouse usage patterns out of this analysis.

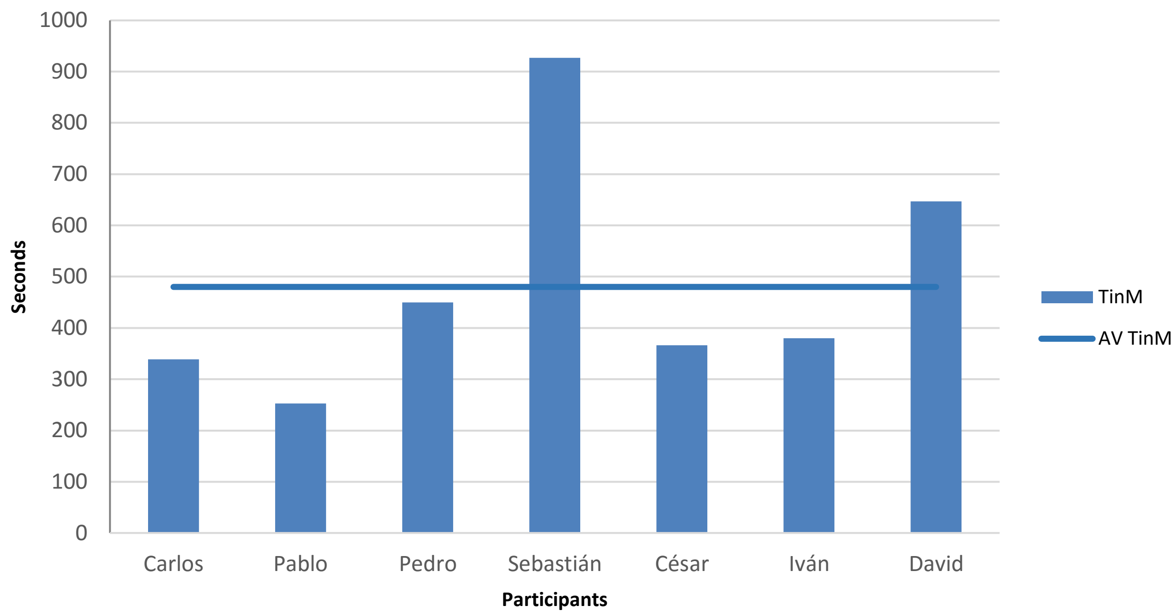


Figure 20. Time in mouse (TinM) (in seconds) per participant in the control group and average (AV).

It is not surprising that these numbers are much lower for the test group. Figure 21 below shows the total time in mouse per participant in the test group, and average. Note that the figure is still labelled as "time in mouse" (since Inputlog does not really know the difference between clicking with a mouse or tapping directly on the screen), but it actually represents all the actions that may be done on the touch-screen with one's fingers or with a stylus, replacing mouse actions such as clicking or scrolling up and down with the mouse wheel. In fact, because touch and stylus are direct input modes (see footnote 29 on page 42), there is not really a "moving-the-mouse-around" action; users may tap on the screen directly where they want to "click", and scroll up and down (or even zoom in or zoom out) using the stylus and/or their finger(s). From this group, it is remarkable that Andrés (T2) spent the most time performing those "mouse-like" actions, from more than twice to almost 20 times more time than other participants in this group. Note also that his time in active writing mode was also the second highest in his group (487s), and that his grade was nonetheless above the average (6.3).

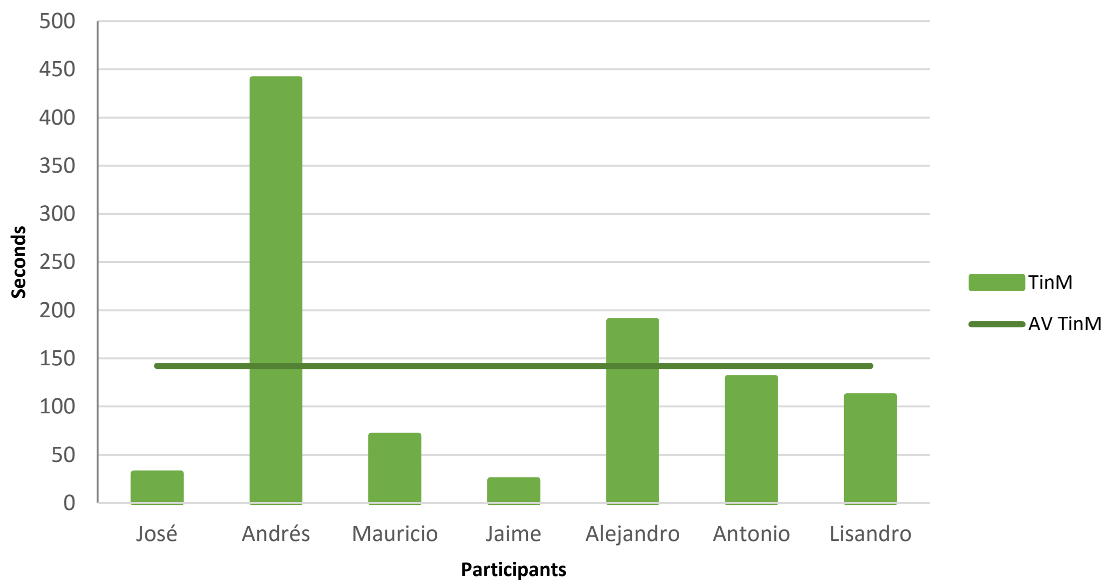


Figure 21. Time in mouse (TinM) (in seconds) per participant in the test group and average (AV).

Now, let us add window statistics to this time-in-mode analysis. Window statistics are measures of user activity in specific windows and applications, for instance, how much time was spent by the user in a particular application or website. Figure 22 below provides window statistics for the control group, and averages. Window statistics presented in this thesis are limited to time measures, and the focus is put on the target-text and source-text windows in MS Word, and the Terminotix toolbar window, which is the window that pops up when a query is launched via a voice or a keyboard shortcut using the Terminotix toolbar add-on (see Figure 10 on page 103).

The time spent in other windows or applications have been gathered together, given the scope of this thesis. Other windows and applications include Google Chrome (Google searches and navigation in other websites), Taskbar in Windows, Task Switching in Windows, the Save As window in MS Word, Dragon NaturallySpeaking's dictation box and spelling window, and other applications unidentified by Inputlog.[73]

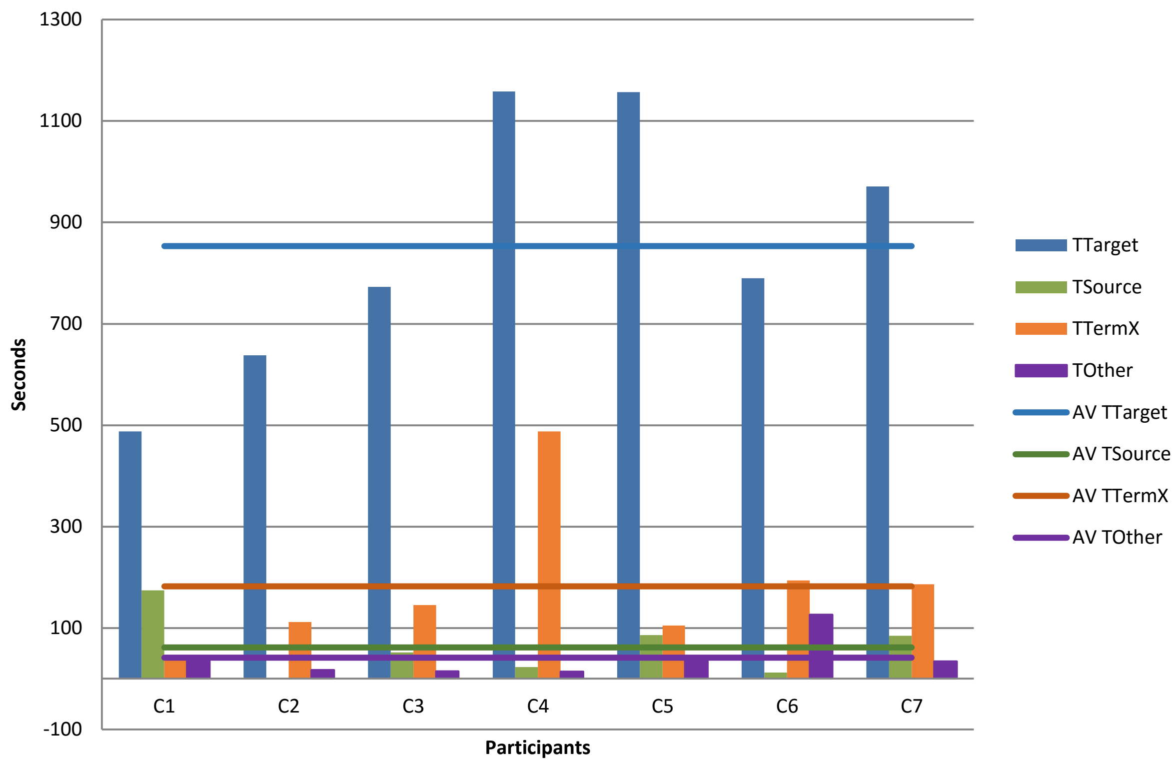


Figure 22. Window statistics for the control group, and averages (AV). Time in target (TTarget), time in source (TSource), time in Terminotix window (TTermX), time in other (TOther), all in seconds.

[73] A reminder to the reader that input data collection was only possible with the PC, and not with the tablet, which was also used by all participants to search information, however minimal that interaction was. Ideally, window statistics from the PC should have been complemented with those from the tablet.

The combined analysis of these task-time data and screen recordings points to different user behaviour. For instance, participant Carlos (C1), who had the lowest total translation times (750 s.), was also the one who spent the least time in the target text window (488 s.) and the Terminotix window (45 s.), but spent the most time in the source text window (174 s.), more than twice the time spent in the same window by any of the other six participants. An observation of the screen recording reveals that this participant kept the source and target text documents open and switched back and forward as necessary as he translated the text. Participant Pablo (C2), for his part, spent less than one second (0.8 s.) in the source text window. A quick glance at the screen recordings revealed that this participant rapidly copy-pasted the entire source document into the target text window to avoid having to constantly switch between the two windows. He simply overwrote the source text in the target text document and did not go back to the source text document.

In addition, participants Sebastián (C4) and César (C5), who spent the most time in the target text window (1158 s. and 1157 s. respectively) are also those with the highest total translation times in this group (1410 s. and 1700 s. respectively). Sebastián (C4) also copy-pasted the source text into the target text document, but still spent 23 seconds in the source text document because he appears to have forgotten that he had copied it, and went back to the source text window after searching the first term. In addition, observing screen recordings may also help in explaining why participants spend so much time in the target text window. Again, in the case of Sebastián (C4), when overwriting the source text in the target text window, a considerable amount of time is spent deleting the source text, sometimes one backspace at a time. Participant César (C5) also overwrote the source text in the target text window. By observing the screen recording, it can be seen that before copy-pasting the document, this participant did spend some time in the source text window (86 s.), reading first, but also trying to figure out how to split the window so he could have the source text in sight while overwriting the other one.

As far as the Terminotix toolbar window is concerned, it was the participant with the highest total translation time (Sebastián [C4], with 1700 s.) who also spent the most time in this window (488 s.). This is more than double the time spent by anyone else in the group in the same window. The observation of the screen recordings reveals that while this participant tried to search most of his information using the toolbar add-on, the information he found was not relevant sometimes, or he found it difficult to navigate in the information sources.

To conclude, no particular trend can be noticed with the combined analysis of these window statistics and screen recordings. Each participant presents different behaviour and different skills, as was also noted in the pre-control task. Another explanation is that they all have different profiles (age, level of experience, varying skills with technologies), and that the outcome of the translation task is not supposed to match one optimal solution (see further discussion in the following subsection).

If participant Carlos (C1) had the lowest translation time, this could be partly explained by the fact that he did not overwrite the source text in the target text window. Indeed, the latter action seems to be taking a significant amount of the other participants' time. But participant Pablo (C2), who did overwrite the source text in the target text window, had the second lowest total translation time. Only a closer look at each participant's data (including their metadata) can allow the researcher to formulate adequate hypotheses about their behaviour. For instance, can we say, based on the data we have presented so far, that participant Pablo (C2) was the most successful of all performing the experiment? We do know that he had the second lowest total translation times, the second-highest grade, the least time spent with the mouse, and perfect transcription accuracy in the pre-control task. We also know that he copy-pasted the source text into the target text window very quickly and therefore had the source text at sight while translating. The screen recording revealed that he is very skillful with the keyboard, in particular to select and quickly delete big portions of text. But the hypothesis that he was the most successful can only be further supported by other data, including metadata (which cannot be discussed in this thesis because of anonymity reasons).

Let us now discuss window statistics for the test group. Different behaviour can be observed in this sample. Figure 23 below provides window statistics for the test group, and averages. A quick glance at the graph reveals that participants spent much less time in the source and target text windows than participants in the control group did. They also spent much more time in the Terminotix toolbar window and in other windows than participants in the control group did. In particular, participants Andrés (T2) and Lisandro (T7), who had the highest total translation times, are also the ones who spent the most time in the Terminotix toolbar window. More precisely, Andrés (T2)'s data reveal that out of 2280 seconds (total translation time), a staggering 1388 seconds were spent in the Terminotix toolbar window browsing information (i.e., searching for translation solutions or term equivalents in Termium Plus and/or Linguee), which also explains the considerably higher amount of time spent "with the mouse" (which in this case of the touch-screen means scrolling up and down with the fingertip or the stylus).

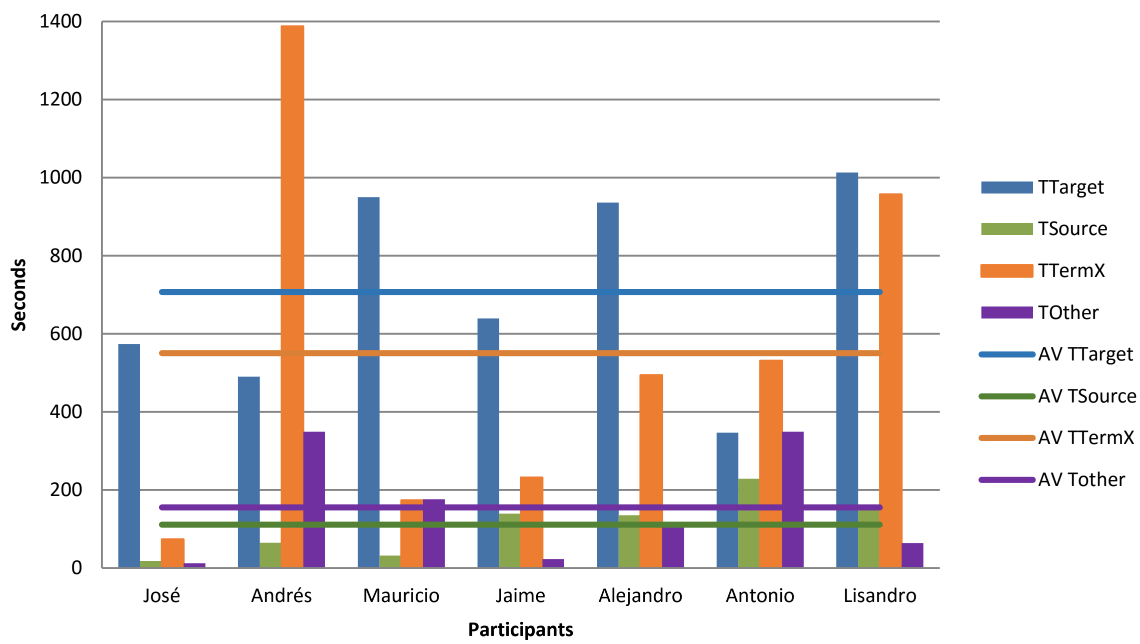


Figure 23. Window statistics for test group, and averages (AV). Time in target (TTarget), time in source (TSource), time in Terminotix window (TTermX), time in other (TOther), all in seconds.

As noted above, the time spent in the source and target text windows was lower for the test group than it was for the control group. It was the time spent in the Terminotix toolbar and other windows that made the whole difference in terms of total translation times (the Terminotix window in particular). The reasons for this are to be investigated in future work, by closely examining screen and video recordings.

Let us now analyze another objective measure, that is, usage patterns, which, combined with task time measures, allow to formulate stronger hypotheses about the TX and the quality-in-use of the environments investigated.

6.2.2.4. Usage Patterns

The focus of the analysis in this subsection will be on keyboard usage patterns, while demonstrating why it is necessary to find relationships between these measures and other types of objective measures, and with subjective measures.

It is widely known in the translation sphere that there are as many (good) versions of a translation as there are (good) translators. Unlike the pre-control task, in which participants were all expected to copy the same text, and therefore the outcome of the task was expected to be identical for all participants, the result of the experimental task gives us 14 different versions of varying length (and quality, as already discussed). Combined with a participant's typing speed, the length of the product of the interaction may also have an impact on the total task times (i.e., the longer the target text the longer it takes to type it). This is another reason why task times alone cannot be considered as indicators of a participant's efficiency and of the quality-in-use of a system in TCI research, and that is because the product of the interaction is not expected to match *one* ideal or optimal solution.

As has been the case so far in this analysis, the length of texts is measured in terms of the number of characters including spaces, instead of the number of words. The Inputlog software provides data on the number of characters of a text (final product), but also on the number of characters typed during the process to arrive at that product (meaning that all missed characters and deletions during the process are also counted). Although no comparison between participants would be relevant in this case, it is possible nonetheless to compare a participant's overall technical effort with the number of characters of their own final product. Figures 24 and 25 below illustrate the number of characters including spaces in the process vs. the product per participant in the control and test groups respectively, and averages. From both figures, it can be concluded that the length of the 14 translations (products) are relatively comparable (between 725 and 892 characters; 819 characters on average). Figure 24 reveals, however, that the technical effort needed by participants in the control group to produce a translation using a physical keyboard exceeds the total number of characters of the product by nearly 20% on average – a considerable percentage to take into account when measuring task times in active writing mode (see Figure 18 and Figure 19 on page 133). As far as the test group is concerned, what Figure 25 reveals is that the technical effort (in this case, the total amount of characters typed *with Windows 8's virtual keyboard*) is at least 1/3 of the total number of characters of the product (in the case of Alejandro [T5]) and can go down to 1/30 (in the case of Antonio [T6], who only hit the virtual keyboard 27 times to produce a text that is 830 characters long – the rest was done using his voice).

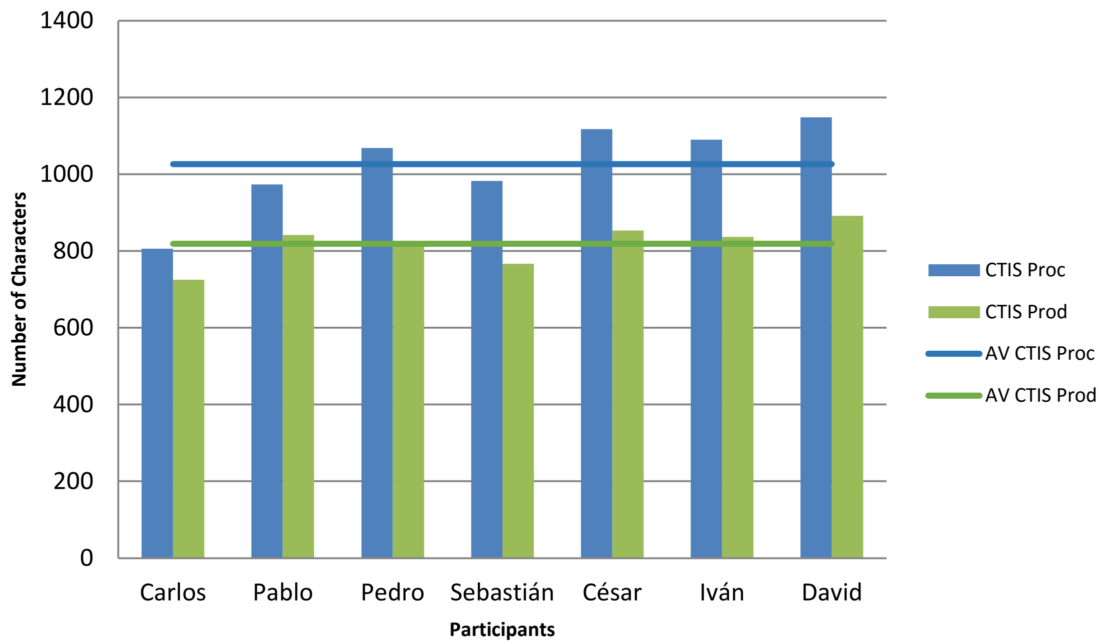


Figure 24. Characters typed including spaces in the process (CTIS Proc) vs. product (CTIS Prod) per participant in the control group and average (AV).

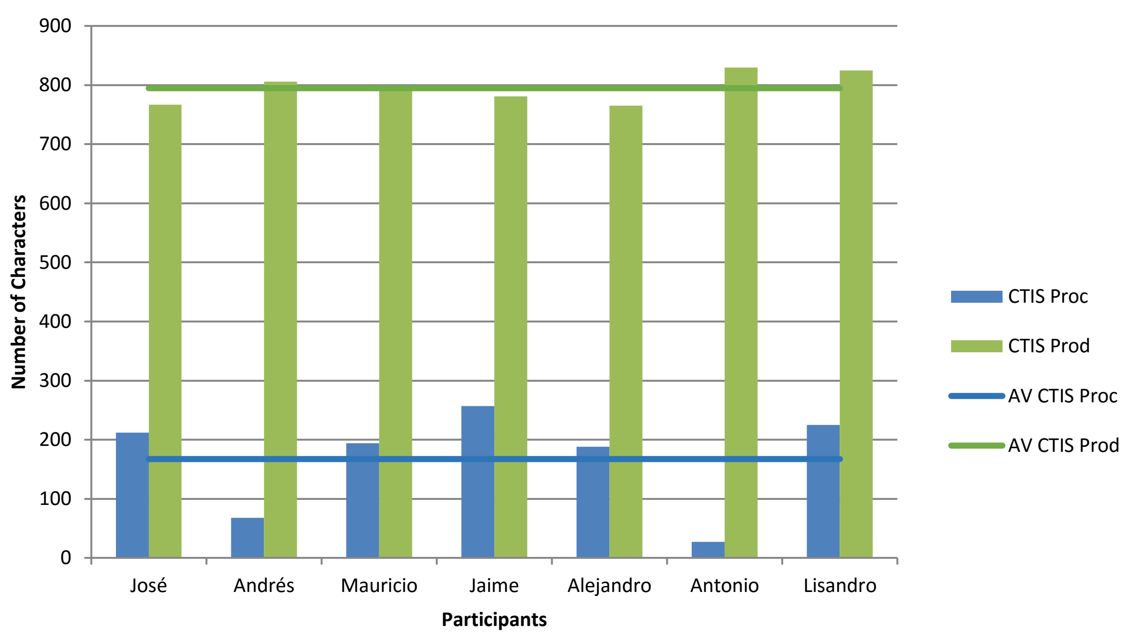


Figure 25. Characters typed including spaces in the process (CTIS Proc) vs. product (CTIS Prod) per participant in the test group and average (AV).

Likewise, when looking at texts at the character level, there is another factor, often under-investigated in TPR and TCI, of considerable importance when measuring the quality-in-use of a particular interface or environment, and overall task performance: the number of typed spaces. A text is composed of both words *and* spaces. It is rarely acknowledged, however, that there are almost as many spaces in a text as there are words. In addition, as the results of this experiment (and other pilot experiments I carried out) show, the number of typed spaces represents around 10% of the total number of keystrokes. Again, this percentage becomes significant when considering the technical effort and the total time needed to type a text using a physical keyboard.[74]

Figures 26 and 27, respectively, display a comparison of the number of spaces typed during the process versus the number of spaces in the final product per participant in the control and test groups respectively. Needless to remind the reader that, in the case of the test group, spaces were typed on the Windows 8's virtual keyboard. Antonio (T6) from the test group, for instance, only needed to hit the virtual keyboard's space bar 4 times to produce an 830-character text. David (C7), from the control group, hit the space bar 167 times to produce an 819-character text. In short,

[74] Let us note that certain efforts in touch-screen-based keyboard application development have acknowledged the significance of typed spaces in terms of technical effort and time, as discussed here, and have strived to remove the need to type spaces in between words. For instance, the Swype keyboard in the tablet used in this experiment automatically inserts spaces after each correctly swiped word, or Windows 8's virtual keyboard, which inserts a space when a user selects a word from the autosuggested words.

nearly 20% of the technical effort by David (C7) to type his translation was hitting the physical keyboard's space bar.

When comparing the number of spaces typed during the process and the number of spaces in the final product, the trend is not surprisingly similar to that with the total number of characters typed (shown in figures 24 and 25 above). The total number of spaces typed in the process exceeds the number of spaces of the product in the control group. As for the target group, the number can be 1/3 down to 1/30.

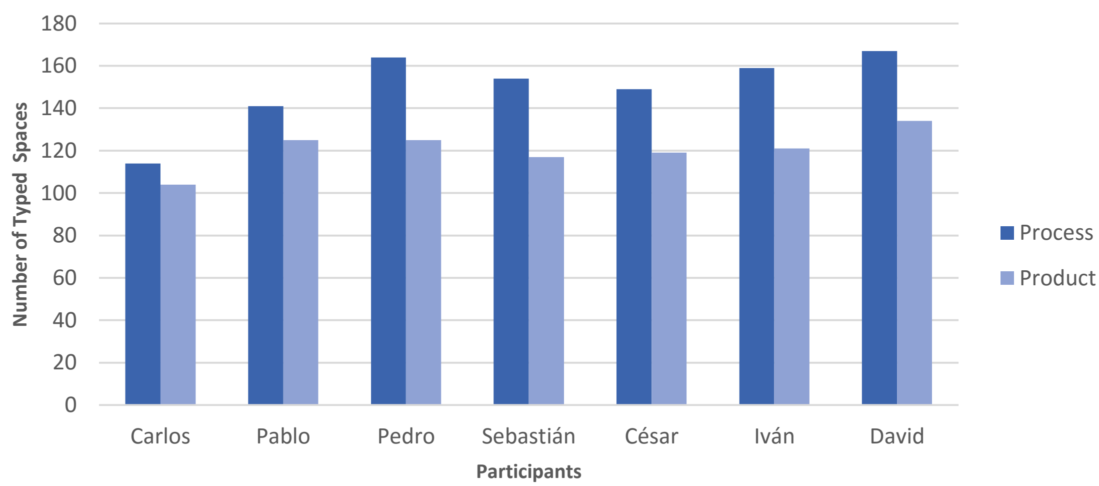


Figure 26. Comparison of spaces typed in the process vs. spaces in product per participant in the control group.

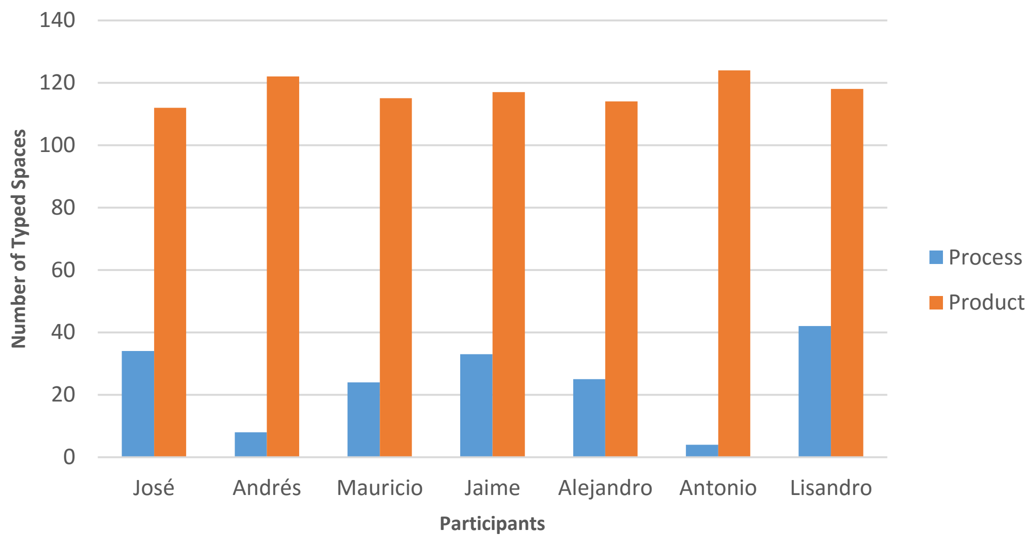


Figure 27. Comparison of spaces typed in the process vs. spaces in product per participant in the test group.

Another relevant usage pattern measure that can be obtained with Inputlog is the number of switches between keyboard and mouse, which, in practice, may have an impact on the overall technical and physical effort and task times (i.e., constantly switching from keyboard to mouse may be a repetitive and tiring action, and may take a considerable amount of time during the interaction process). Figure 28 shows the number of switches between keyboard and mouse per participant in the control group, and average, as logged by Inputlog. Not surprisingly, Sebastián (C4), who had the highest total translation time in the control group and who also spent the most time with the mouse, also had the most switches between keyboard and mouse. It can safely be concluded that this participant is not very comfortable or skilled with traditional input devices and that these devices somehow impact negatively his performance in terms of efficiency.

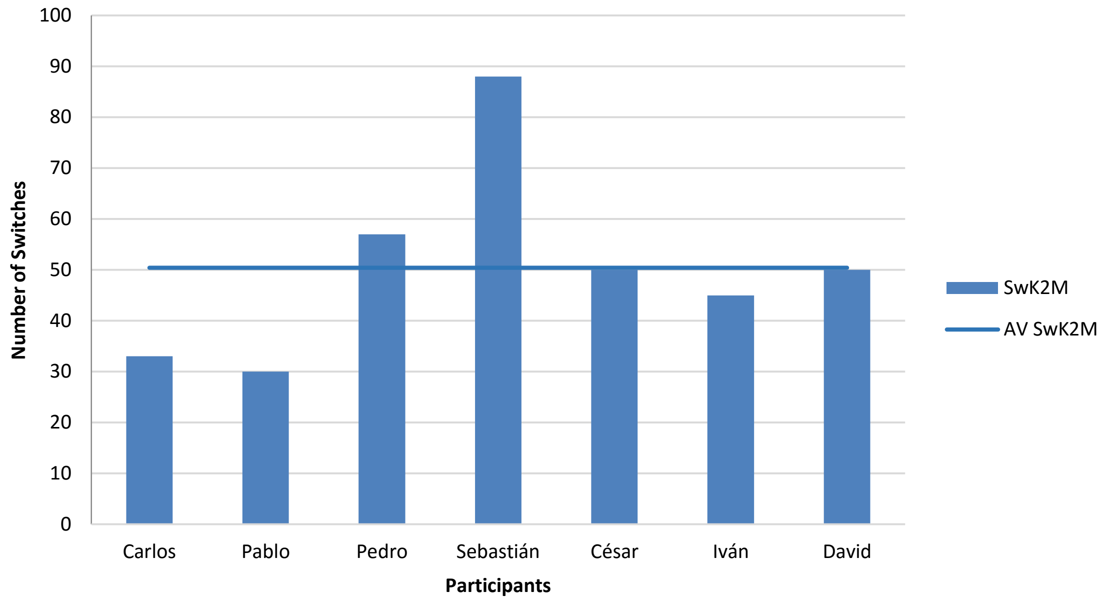


Figure 28. Switches keyboard to mouse (SwK2M) per participant in the control group and average (AV).

Now, does the multimodal environment reduce the need to constantly switch between input modes? Figure 29 below shows the number of switches between keyboard and "mouse"[75] per participant in the test group, and the average. The voice-and-touch capabilities of the multimodal environment seem to reduce to a minimum the number of switches needed between input modes. It can be observed nevertheless that participants Andrés (T2) and Lisandro (T7), who also had the highest translation times overall, are the ones in this group who did the most switches between typing with the virtual keyboard and tapping somewhere else on the screen.

[75] Again, Inputlog is not measuring mouse activity here but mouse-like actions such as tapping, performed outside of the virtual keyboard application.

Still, the average number of switches for this group remains less than half of the average number of switches for the control group.

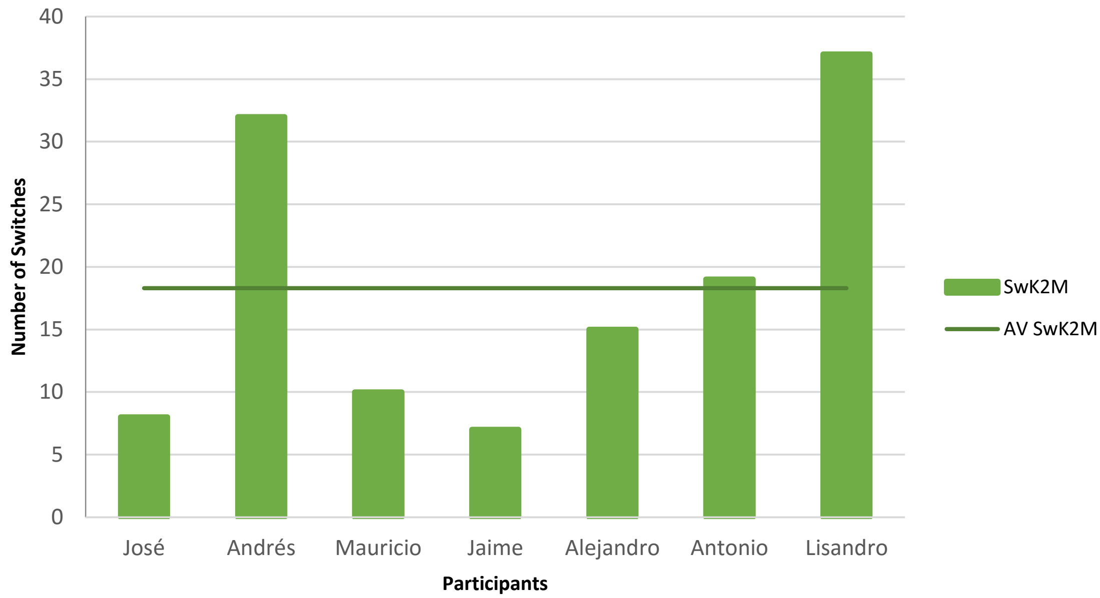


Figure 29. Switches keyboard to mouse (SwK2M) per participant in the test group and average (AV).

To conclude the usage patterns analysis, let us include other kinds of data provided by the Inputlog software: the number of times users go from a window or application to another, or number of transitions between windows. Figure 30 below shows the number of transitions for the control group, and average. A number of links can be made between these and other data. For instance, Carlos (C1), from the control group, had the most transitions between windows (98 transitions), but this is because he kept the source and target text documents separate and needed to go back and forth repeatedly during the process. Pablo (C2), for his part, who is the same participant that we hypothesized above was the most successful performing the experiment, had the least transitions between windows (22 transitions), the least switches between keyboard and mouse, spent the lowest time with the mouse, and had the second lowest total translation time, and the second-highest grade in his group.

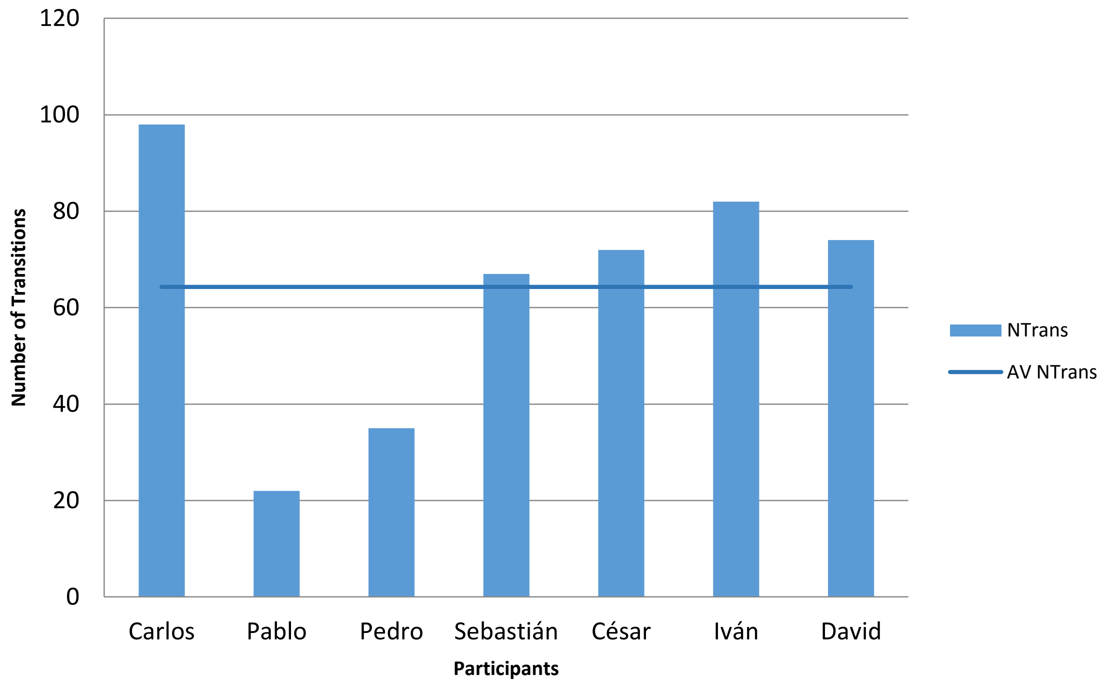


Figure 30. Number of transitions (NTrans) between windows for the control group and average (AV).

Not surprisingly, the number of transitions is higher for the control group than for the test group. Figure 31 below shows the number of transitions for the test group, and average. In particular, it can be noticed that the number of transitions between windows seems to be linked with total translation task times: the two participants with the shortest translation times also have the lowest number of transitions between windows; the two participants with the longest translation times also have the highest number of transitions between windows.

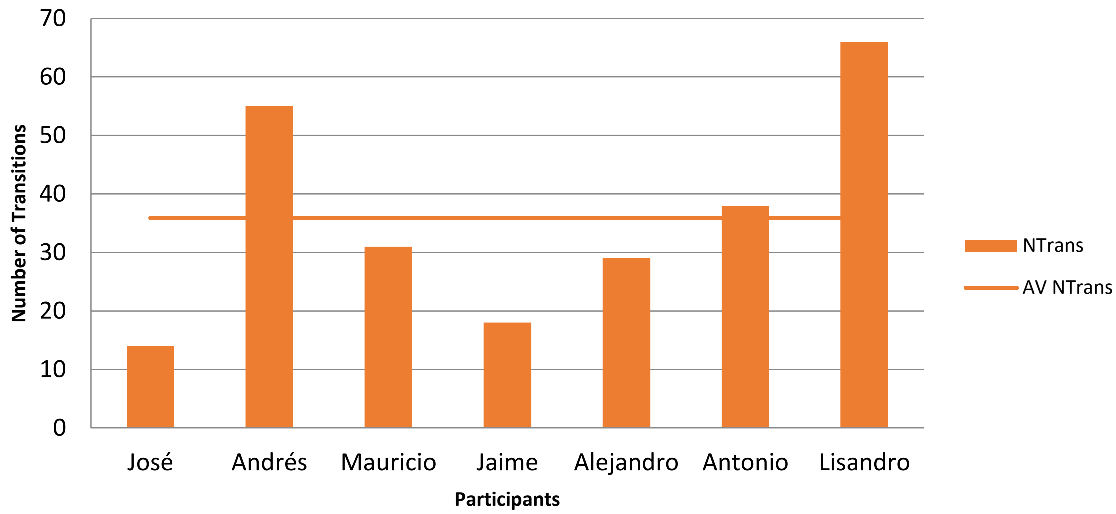


Figure 31. Number of transitions (NTrans) between windows for the test group and average (AV).

6.2.3. Concluding Remarks on Quantitative Data Analysis

In this section, I presented objective outcome and interaction process measures. Looking at averages alone, up to this point, the following can be concluded: when considering total translation times, participants in the control group performed better than those in the test group (1138 s. vs. 1517 s.); they took 25% less time to complete the task with the KM environment than participants with the experimental ITD environment. However, only a minor difference in terms of objective quality could be observed. Furthermore, the different input process measures provided by Inputlog indicate that the technical effort needed to produce a translation with a traditional KM environment always exceeds the technical effort needed with the prototypical ITD environment (in terms of number of characters typed, which includes spaces and deletions, number of switches between keyboard and mouse, time spent typing in active writing mode, time spent with the mouse, number of switches between windows, etc.).

Now, are these data enough to affirm that the prototypical ITD environment provides (or not) a better TX than the KM environment? While most of the objective measures that have been discussed so far seem to point towards a better quality-in-use of the experimental environment, the need to complement such data with qualitative data seems to be necessary. As Swallow et al. (2005) put it, "qualitative data provides a richness and detail that may be absent from quantitative measures."

Indeed, the combined analysis of both objective and subjective measures should offer a bigger picture of the TX and of the quality-in-use of the different environments tested. In addition, the qualitative data from the experiment's interview should also provide information about the participants' current workflow and workstation.

Let us now look at the qualitative data in more detail.

6.3. Qualitative Data Analysis

In this section, qualitative data from the interviews are presented and discussed. Qualitative data analysis constitutes the second step towards confirming or falsifying my hypothesis, that is, whether or not the prototypical ITD environment consistently provides translators with a better TX than the KM environment. Various measures will be presented and discussed here: users' perception and assessment of the outcome, subjectively experienced duration, users' perception of interaction and task difficulty, ergonomics, users' sense of success, and satisfaction.

During the interviews, many topics were broached, from the current situation at work to the current advantages and drawbacks of the devices and interaction modes tested. As I mentioned in the previous chapter, the interview questions (see appendix B) were designed to look beyond the experimental tasks and provide a bigger picture of the TX with the prototypical ITD environment versus a traditional KM environment. Thus, most of the participants' answers were retrospective, based on their current experiences at work, or on the idea of integrating into their workflow devices and software similar to the ones used in the experiment.

I would also like to remind the reader that I needed to reformulate the interview answers for two main reasons: (1) when answering an interview question, a person may not organize their thoughts coherently before formulating an answer, so answers were reformulated in a syntactically and grammatically correct language, while taking special care not to leave out any important information; and (2) as per research ethics regulations, anything that may reveal the participants identity or that of their organization needed to be removed. The data is also presented in an aggregate format by category and with no reference to the participants (not even to their pseudonyms) in an effort also to maximize anonymity.

In addition, most of the answers are paraphrased in English or translated into English for the purposes of the present thesis. In some occasions, the exact wording in French is provided in a footnote where the translation may be challenging or might not fully reflect the nuances of the original phrasing.

6.3.1. Current Work Environment

I included in the interviews a number of questions regarding the participants' current situation at work (see interview guides in appendix B). The present subsection gathers these data and highlights both positive and negative attitudes towards their current workflow and workstation.

6.3.1.1. Workflow and Workstation Description

Participants were asked to describe their current workflow and workstation (software and hardware used). A common translation workstation, as far as hardware is concerned, is the typical KM one-screen desktop computer. However, it can be noticed in the data that other configurations have been proposed to translators in an effort to increase their productivity and to prevent illnesses.

A recurring subject was the use of two screens, instead of one. The operating system provides the option to extend the view to two monitors, thus users have more space to place different windows and applications. Most translators in our sample reported that they use two screens; however, not all of them use them the same way or for the same purposes. For instance, some participants said that they place the source text on the screen at the right side and work (translate) on the left side: "It's more natural for me"[76], says one translator. In contrast, other participants said that they place both target and source texts on one screen, and keep the second screen for browsing information. Interestingly enough, one participant said that he uses both approaches: sometimes the French text on one screen and the English on the other screen, and

[76] *C'est plus naturel pour moi.*

sometimes the two texts on one screen and the tools on the other. As far as the information-browsing screen is concerned, some participants said that they cannot have more than two or three windows or tabs open because if they do, the system slows down, or produces an error and freezes.

The keyboard and the mouse also came into play. Some participants described different ways to use these input devices: for example, setting the keyboard in the middle of the two screens, or using shortcuts “to avoid having to extensively use the mouse”[77]. A couple of participants also mentioned the use of special devices such as “ergonomic” keyboards and mice. They said that such an environment had been proposed by a professional work ergonomist. One of them described the ergonomic keyboard as a keyboard that is not straight but rather split in two halves, positioned in an angle, so that the arms rest in an “optimal position”.[78] The ergonomic mouse was described as a big mouse that is adapted to the size of his hand. In addition, he explained that he uses a separate keyboard for numbers (i.e. a numeric keypad); that his chair is ergonomic and adapted to his height; and that the two screens are placed in an ergonomic way, meaning that they are not placed forming a straight line but rather in an angle.

Another element of the current workflow and workstation to be kept for the purposes of this analysis is the use of pen and paper. Translators still use physical pen and paper for different purposes. For instance, one translator explained that he takes

---

[77] *J’utilise énormément les raccourcis clavier pour éviter d’avoir à utiliser la souris.*

[78] *Position optimale.*

notes on a piece of paper, "for words that appear over and over"[79]. Another translator said that he prints the source text, underlines the terms he needs to search, takes notes on paper and then starts translating. In contrast, a couple of translators said that they print the semi-final draft of the target text at the revision step of the process: "It is easier to see any mistakes,"[80] said one of them.

To add to this current work environment description, I gathered all mentions of different software and applications. This gives a researcher an idea of the tools in a translator's toolbox, and what types of tools are used more than others. As observed in Figure 32 below, translators in the sample use an array of tools, from dictionaries to language guides, from spelling and grammar checkers to term banks and TMSs.

[79] *Je me fais des mémos à côté pour les mots qui reviennent régulièrement.*
[80] *Les fautes ressortent beaucoup plus.*

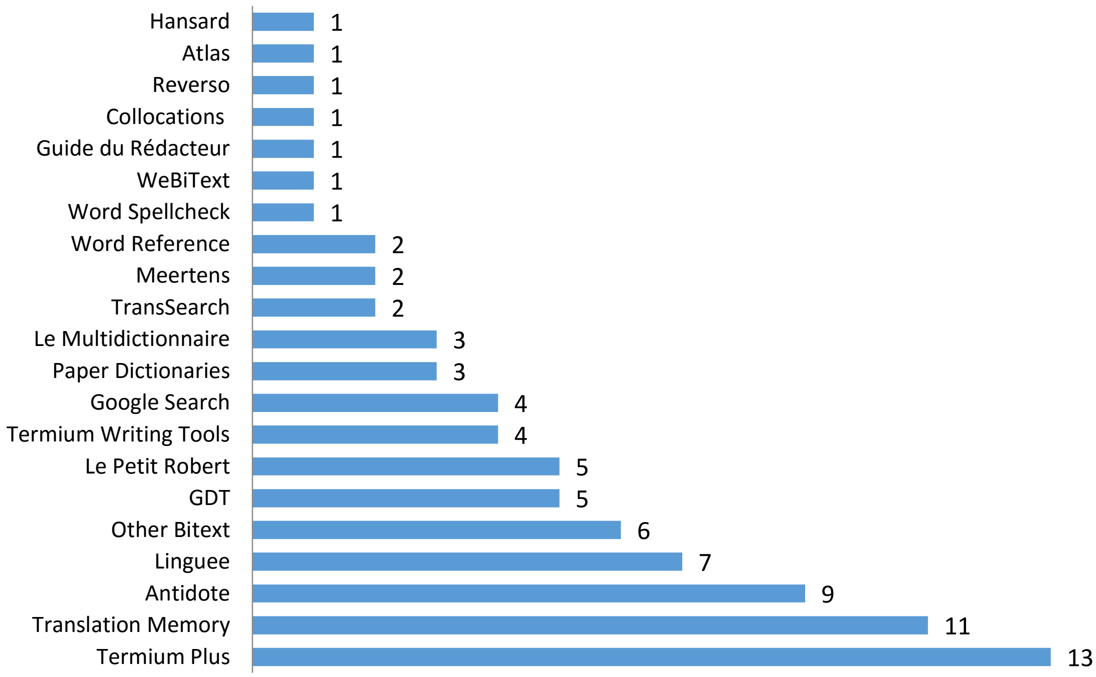


Figure 32. Tools used in current work environment, and number of participants mentioning each tool.

In Zapata (2016), I report on a pilot study investigating the use of various information resources by professional translators. Based on my experience with that study, I recognise that a more detailed analysis on the current use of tools (by closely examining screen recordings for instance), and the translators' preferences, is beyond the scope of this work.[81] Such an investigation will be necessary, however, in future work aiming at improving information tools and integrating them to translation environments, including an ITD environment. In the following subsection, however, I

[81] Other studies investigating translators' interactions with resources include Désilet et al. (2009), Ehrensberger-Dow and Massey (2014), Enríquez Raído (2014) and Volanen (2015).

do present some positive and negative attitudes towards the current tools and workstation.

6.3.1.2. Satisfaction with Current Tools and Workstation

In one of the interview questions, I asked the participants to rate on a scale from 1 to 10 their level of satisfaction with the current tools and environment used at work. Their answers ranged from 5 to 9 out of 10 (7.3 on average). This may indicate a high level of satisfaction. Nevertheless, a close examination of the participants' comments that followed the question better informs on the TX with the current work environment. One participant for instance said that he would rate his level of satisfaction 3 out of 10 for some aspects, and 9 for some others. Attitudes were both positive and negative, but negative comments were predominant. I acknowledge that even though quantifying these attitudes would have been ideal for the rigour of my analysis, it was not possible to do so because some comments needed to be excluded to ensure the participants' anonymity and that of their organization.

Note also that some of the topics discussed in this section are not linked specifically with the tools and techniques explored in this study, for instance, the use of TMSs, for which no quantitative data were collected. However, this exposition may add to the researcher's reflections on the TX with different types of tools and inform future work in that area.

6.3.1.2.1. Positive Attitudes

One translator was rather enthusiastic about the current work environment. He said that he is "pretty satisfied"[82], that he can produce a text that is "good enough"[83] in relatively little time, and that the tools that he has allow him to do that. Other positive comments by other participants include that the tools' performance has significantly improved over the past few years, that they usually find what they are looking for, and that "some of the things that come from the translation memory are good"[84]. Other positive statements include: "Overall, everything works very well"[85], and "I'm satisfied. I don't have any problem"[86]. One translator remarked that he is satisfied with his current workstation, "not because it is the best out there, but because I'm used to it"[87]. In the same vein, another participant stated that "using the keyboard is like wearing your old slippers;"[88] it is a matter of being used to certain tools which makes it difficult to want anything else. Now, as we sank more deeply into the conversations, participants raised a number of concerns; negative comments dominated the discussions.

6.3.1.2.2. Negative Attitudes

A major concern with the current work environment is the inability to meet deadlines or not being able to deliver high-quality texts in some instances due to limited time to

---

[82] *Assez satisfait.*
[83] *D'assez bonne qualité.*
[84] *Il y a de bonnes choses qui viennent de la mémoire de traduction.*
[85] *Dans l'ensemble, ça va super bien.*
[86] *Je suis satisfait. Je n'ai pas de problème.*
[87] *Non pas parce que c'est le meilleur, mais parce que je suis habitué.*
[88] *Le clavier, c'est comme de vieilles pantoufles.*

perform all tasks. One translator stated that given that deadlines are tighter and tighter, there is a moment where you no longer have the time to stop and do any research, and thus deliver a good-quality text.

Although CAT tools' main purpose is to increase efficiency (TMSs for instance, widely used by translators in our sample), some translators appeared to be unsatisfied with their use (see also Leblanc's (2013) study). One translator stated specifically that he does not like translating directly in the TMS because he does not have enough context, so he gets the impression that there will be something missing. Another participant said that he would like it if he was given the choice to use one TMS over another; he does not like that the tool was just imposed on him. One translator argued that, at work, he and his colleagues are obliged to use a TMS because the expected productivity depends on it; they cannot just decide to translate from scratch in a Word document. One of the participants elaborated further on his discontentment with the TMS: "It is mostly the translation memory that I don't like"[89] because it automatically replaces the paragraphs and sentences that have been already translated, and sometimes, he said, "it doesn't help at all"[90]. He explains that there are paragraphs that are mistranslated, and that taking into account the time to read, realize that it is not good, and re-translate, it is easier to translate from scratch. In the case of long texts, using "polluted"[91] TMSs becomes "a big waste of time"[92]. In addition, two participants

[89] *C'est surtout la mémoire que je n'aime pas.*
[90] *Ça n'améliore rien.*
[91] *Polluées.*
[92] *Une grosse perte de temps.*

referred to the TMS they use as not being the most "user-friendly"[93]. Lastly, the segmented nature of the work with a TMS was also mentioned: "I wasn't very happy with it […] I didn't like working segment by segment [when I used a TMS many years ago]."[94]

A number of concerns are related to the computer's power and capacity. One translator, for instance, complained about recurrent system breakdowns or about the operating system being too slow. He says that "we become a little too dependent"[95] on the speed and robustness of computers. Another participant seemed to think that there is incompatibility between the different applications, including CAT tools, which cause the system to freeze or break down. One participant was concerned, in particular, about the combined use of two computationally heavy applications such as a TMS and a VR system.

A general concern is also the fact that there are too many tools and resources which work separately, not integrated: "There are too many tools all over the place,"[96] said one participant; another one said that the biggest problem right now is that there are too many tools and too many windows to keep open. In the web browser, he explained, he has five or six tabs open, and he does not have any other choice than to switch from one to the next.

---

[93] *Convivial.*

[94] *Je n'en étais pas très satisfait, je n'aimais pas travailler en blocs.*

[95] *On est un peu à la merci de la force.*

[96] *Il y a trop d'outils disparates.*

Other concerns relate to the hardware, in particular the screen and the mouse. One participant said that, although he uses two large screens, he finds that to be "useless"[97]. He continued saying that he feels "bombarded by the brightness of all that"[98]. Furthermore, the mouse seems to be a source of frustration for many participants. One of them complained about having to do "just too many clicks"[99] and said that one of the things that bothers him the most is having to reach for the mouse to search for something: "I think that I waste way too much time clicking, always clicking,"[100] said the same person.

Clicking, as well as typing, indeed takes a considerable amount of a translator's time. But the ergonomics of a device or working environment also has an impact on the TX, as will be discussed below.

6.3.1.3. Ergonomics

The ergonomics of the working environment was a recurring topic during the interviews. Only one participant appeared to be, without any hesitation, satisfied with the current work environment from an ergonomic viewpoint: "I have never been physically tired from work",[101] he said, when the chair and the keyboard are in a good position. But the rest of participants do seem preoccupied about physical discomfort, feeling pain or developing illnesses due to working long hours sitting in front of a desktop computer: "For me, that is the most difficult",[102] said one of them.

[97] *J'ai deux énormes écrans, mais je trouve ça même inutile.*
[98] *Bombardé par la luminosité de tout ça.*
[99] *Énormément de clics.*
[100] *Je pense que je perds énormément de temps à cliquer; toujours cliquer, cliquer.*
[101] *Je n'ai jamais ressenti de la fatigue au travail.*
[102] *Pour moi, c'est ça le plus difficile.*

One participant stated that spending a day working with the keyboard and mouse can be painful; that he can feel it in his wrists by the end of the day. Another translator said that a few years ago, he was suffering from stress and wrist pains. He argued that typing for a long time in the same position without moving anything but the fingers was putting a pressure on his wrists, and explained that his doctor ordered him to obtain an ergonomic assessment because “typing all day long is not good”.[103] He concludes affirming that, in his opinion, every translator has the same problems: pains in wrists, elbows, fingers, arms and back. One current solution, for him, is to vary his tasks so that he is not actually making the same movements all the time. This is one of the reasons why he prints his translation to revise on paper because that “makes me use my hands and my eyes differently”. One participant mentioned having been advised by his physiotherapist to take breaks at work, and stand up from time to time, but that he cannot do that with pending deadlines or knowing that he is lagging behind the expected daily productivity.

One of the participants who mentioned the use of an ergonomic keyboard and mouse at work explained that he wanted to obtain an ergonomic assessment at a young age, because he does not know the long-term effects of working on a traditional desktop PC. He said that he is aware that translators are very likely to develop tendinitis because they spend 95% of their time typing on a computer. So, he continued, by having an ergonomic workstation, he thinks to be postponing or preventing illnesses and injuries. He acknowledges that he calls his workstation “ergonomic” and

[103] *Taper la journée longue n’est pas bon.*

"optimal" in the context that he is in, because he does not have the choice. This being said, "a workstation like [the multimodal ITD environment] would truly be a whole new level"[104], he affirmed.

With this in mind, let us now move onto the qualitative data that provides insight into the TX with the experimental environment.

6.3.2. Experimental Environment

A number of questions in the interview aimed to investigate the participants' perception of and responses to the use or anticipated use of an ITD environment (see interview guides in appendix B). The present subsection gathers some of these data together and highlights both positive and negative attitudes towards emerging input modes, and emerging computational devices such as touch-screen computers and tablets.

In order to make a smooth transition from the previous section, let us first examine elements from the discussion that regard ergonomics.

6.3.2.1. Ergonomics

Many advantages are attributed to ITD from an ergonomics viewpoint. For translators in our sample, the greatest advantage of ITD appears to be the possibility to work standing up and prevent illnesses and injuries: "Eventually, if ever I have more pain in my wrists than I have right now, I would certainly love to have a system like this,"[105]

---

[104] *Avec un matériel comme ça, c'est vrai qu'on est dans une classe à part.*
[105] *Éventuellement, si j'avais des problèmes avec les poignets davantage que j'en ai présentement, c'est sûr que j'aimerais avoir un système comme ça.*

stated one translator. Another participant said that it would be "great"[106] if he had the possibility to work standing up from time to time. Another one said that "it feels really good"[107] that ITD gives the translator the possibility to "take a break from the sitting position".[108] Likewise, another participant commented that "it is absolutely true that I could translate standing up while doing other things"[109], and that this would be "a great advantage"[110] for translators. For one of the participants, the advantage of a multimodal ITD environment is that you have different interfaces and devices, he said, which pushes the translator to move: using the stylus, then using the finger, then using the tablet, and so forth. Let us also note that the one participant who did not hesitate to say that he is completely satisfied with the current KM environment, as mentioned in the previous section, did acknowledge later in the interview that it would be an advantage to be able to move and stand up while translating. He said that if he had the two choices, that is, voice and keyboard input, "it would be good".[111]

Now, are there any advantages as far as efficiency and quality are concerned? In other words, can a translation be produced more rapidly or be of better quality with an ITD environment, according to translators? In the following two subsections, let us look at what participants had to say on quality, and then on efficiency, when interacting multimodally.

---

[106] *Génial.*
[107] *Ça fait vraiment du bien.*
[108] *Prendre une pause de la posture*
[109] *C'est absolument vrai que je pourrais traduire debout en train de faire d'autres choses.*
[110] *Un gros avantage.*
[111] *Ça serait bon.*

6.3.2.2. Perception and Assessment of Outcome

In section 6.2.2.1, the assessment of the outcome of the interaction by expert evaluators was presented. Now, the translators' own perception and assessment of the outcome is also important. During the interviews, participants were asked whether they felt that the quality of the product was better or could be better by interacting with a multimodal ITD environment versus a KM environment. Note that participants in the control group could base their answer only on the short introduction to the experimental environment they were given. However, their answers are not being presented separately but rather together with those of the test group.

In this experiment, there was a consistent lack of enthusiasm from one of the participants, with a definite "no" when he was asked if he thought that the quality could be improved with the multimodal environment. Another participant said that quality might not be better but "certainly comparable"[112] to what one could achieve using the KM environment. The rest of participants were rather positive. The one translator who has experience in TD said that "in dictation the quality of the text is better."[113]

Among the reasons why participants think the quality of the outcome was better or could be better with an ITD environment is that, as one participant said, such an environment would allow the translator to reduce the level of stress because more

---

[112] *Certainement, ça s'équivaut.*

[113] *Je pense qu'en dictée la qualité du texte est meilleure.*

words can be translated in less time. Likewise, another participant said to find that "the quality is improved"[114], because it allows the translator to go faster, so he or she has enough time to reread the text and make corrections as necessary. Furthermore, one translator affirmed that the interaction with a voice-enabled environment would result in a text that is more natural because by saying things out loud, you find that there are "more spontaneous ways of saying things"[115].

ITD also forces translators to develop attention to detail. One participant said that "we all make spelling mistakes when typing"[116], so the ITD environment would allow the translator to take some distance from the text, so that he or she can "look at the text with different eyes"[117] when proofreading it. In the same vein, another translator said that he "cannot trust"[118] what the system will have written, so he would take a closer look at the transcription to make sure that the VR program understood well. Similarly, another translator said that, when typing, one may not see typing errors or typos, so this environment forces you to "revise in much more detail"[119].

The fact that a tool is ergonomic is also believed to have a positive impact on the quality of the outcome of the interaction. One participant said: "I think there will be an impact on quality"[120] because the translator would be in an ergonomic

[114] *Ça bonifie la qualité.*
[115] *Des formulations qui sortent plus spontanément.*
[116] *Quand on tape, on a tous le problème des coquilles.*
[117] *D'apporter un autre regard.*
[118] *Je ne peux pas me fier.*
[119] *Réviser beaucoup plus en profondeur.*
[120] *D'après moi, la qualité va s'en ressentir.*

environment designed for optimal concentration, and thus his energy will be kept all day long. There is "no doubt that it would be noticeable in his translation"[121], he said.

Another translator affirmed that ITD can improve the quality of the translation because the translator is obliged[122] to read the full sentence before translating it: "When you start translating without reading the sentence, you get confused,"[123] he argued. Lastly, another reason why ITD is thought to have a positive impact on quality is that there is "more time to do thorough research"[124], to read documentation on the Internet and to go discuss with colleagues about any difficulties with texts: "These are things we no longer have enough time to do because deadlines are very tight,"[125] said one participant.

The multimodal environment appears to have a positive impact on translators' perceived quality of the outcome. Now, do translators have the same positive perception regarding efficiency gains? Participants were asked whether they felt that efficiency was improved or could be improved by interacting with a multimodal ITD environment versus a KM environment. Let us now discuss another indicator of the TX: task times as perceived and experienced by the participants.

[121] *Nécessairement ça transparaît dans sa traduction.*

[122] This assertion may stem from the fact that the researcher explained to participants that in oder to ensure high accuracy rates from the VR system, the user must be able to dictate a sentence from top to bottom instead of pausing one or several times within the sentence while dictating.

[123] *Si on commence à traduire sans lire la phrase, on s'emmêle.*

[124] *Plus de temps pour la recherche en profondeur.*

[125] *Ce sont des choses qu'on a presque plus le temps de faire parce que les échéances sont trop serrées.*

6.3.2.3. Subjectively Experienced Duration

In section 6.2.2.2, it was shown that average total translation times were higher for the experimental group than they were for the control group. But these objective measures alone are not sufficient indicators of the TX and the quality-in-use of a given environment. In other words, it is not because participants in the test group took more time on average to translate one text that the prototypical ITD environment is less efficient than the KM environment. It is thus necessary to consider both subjective and objective measures because this may lead to different conclusions, as it has been argued and demonstrated so far.

As stated earlier, psychologists have long recognized and quantified the difference between objective time and subjectively experienced duration; therefore, using both measures may lead to a better picture of the UX, since "differences between interfaces in objective time may not be found for subjectively experienced duration, and vice versa" (Hornbæk, 2006, p. 92). It is also argued that the importance of task times may have been overestimated in a number of studies since users, designers and owners of an interface may not equally weigh their importance (*ibid*., 80). Indeed, higher task completion times have been observed to be indicators of users' motivation, reflection, and engagement (see for instance Inkpen's study (2001)).

When asked if ITD could have a positive impact on efficiency, one participant answered with a definite “no”. Note that it is the case of the less enthusiastic participant who also replied negatively regarding the quality of the outcome, discussed above.

In general terms, however, enthusiasm towards efficiency was less evident for some of the other participants than it was towards quality. A recurring comment was that for this particular experiment, they did not find that it went faster, because there is a learning curve. In other words, there is a feeling that it is not until the translator has used this environment long enough and has gotten used to it that there will be any efficiency gains; not for first-time users. One participant acknowledged that he is a fast typist; therefore, he did not perceive any efficiency gains within the framework of the experiment. He did admit, nonetheless, that the VR system was making mistakes probably because he was not dictating properly or because the software had not been trained long enough to his voice and writing style. Other participants also believed that there is a learning curve before any efficiency gains can be achieved: "As far as productivity is concerned, I think there will be a dip before going up again. There are many things to be integrated,"[126] said one of them. Another translator said that "in the long run, or even in the near future, I think that voice recognition will be more efficient than typing"[127]. One participant thought that he could probably be "much more efficient"[128] after two, three, four or five days of use. One participant said that he did feel that he translated faster during the experiment, but that he was less

[126] *Côté productivité, je pense que ça ferait un creux avant de remonter. Il y a beaucoup de choses à intégrer.*

[127] *À long terme ou même à moyen terme, je dirais qu'il y aurait des gains en efficacité en utilisant la reconnaissance vocale plutôt qu'en tapant.*

[128] *Beaucoup plus efficace.*

certain about what he was doing. ITD is a bit "destabilizing"[129], he said, because he is not used to it. "But I see the potential."[130]

For some translators, what allows them to do more in less time is the possibility to search for information quickly through voice commands. One participant explained that, at work, he often has to go click on different tools and sometimes it takes too long to find anything, whereas in the experiment, with the multimodal environment, he had only to tap on a word and say "press control eight" or "press control nine" and got a response instantly.

Some of the comments regarding subjectively experienced duration were immediately positive. One translator stated that dictating a translation with VR was "definitely faster than with the keyboard";[131] another one said that that the experimental environment workflow was "much more efficient".[132] One participant affirmed that his efficiency improved by at least 10%, and felt that it could improve even further once he got used to the environment.

Particular aspects of the translation workflow or features of the experimental environment were also mentioned. For instance, one participant said that with this environment, a translator can do a first draft much faster than doing it with the keyboard. Another participant appreciated the efficiency gains that could be obtained from the autosuggestion feature in virtual keyboards. The system's ability to predict

[129] *Déstabilisant.*
[130] *Mais je vois le potentiel.*
[131] *Sans doute plus rapide que le clavier.*
[132] *C'était beaucoup plus efficace.*

the next word allows a translator to save time because there is no need to type full words every time. For example, one participant explained that during the experiment, just by typing "*i*" the system suggested "*imaginez*", and that it was “exactly what I wanted”[133]. He concluded saying that he would very much like to have that feature at work because translators have so many words to type and by the end of the day they feel tired and start making mistakes.

Lastly, having performed the experiment, some participants saw an absolute advantage in ITD efficiency over TT. One participant commented that when actually using VR, “you see how efficient it is. I don’t think it makes sense to continue typing,”[134] and another said that when you actually get to use a system this “efficient”[135], you realize that typing on a computer is “an incredible waste of time”.[136]

Although measuring the ratio between objective times and subjectively experienced duration is difficult with these data alone, it can safely be said that some translators in our sample do perceive increased efficiency with ITD. However, according to them, considerable improvement appears to be dependent on extended use of the environment. In the long run, with training and practice, ITD could likely be more efficient than TT (and other forms of TD).

[133] *Exactement ce que je voulais.*

[134] *Quand on y est confrontés, on voit à quel point c'est efficace. Donc que je pense que c'est illogique de continuer à taper.*

[135] *Performant.*

[136] *Une perte incroyable de temps.*

Now, how difficult was it for participants to perform a translation task using the experimental multimodal environment? In the subsection below, we will discuss users' perception of task difficulty as another indicator of the TX. Again, note that participants in the control group based their answers solely on their experience during the introduction they were given to the test environment.

6.3.2.4. Perception of Task Difficulty

In general terms, it can be said that most participants found the experimental task rather difficult. But positive attitudes were also found: one participant said that the task might have been easier for him because he has experience in interpretation; in other words, he is used to translating orally. However, for most translators in our sample, the task was completely new.

One participant stated that he was "completely lost"[137] during the experiment, and that he knows that he probably would need to develop reflexes that he does not have. He said that his "old reflexes"[138] kept coming back: using the keyboard, or going back to search the same term three or four times (because he has a bad memory, as he said). In the same vein, one participant said that he has a tendency to go back to his old working habits, mostly when revising. He said that he felt like typing his corrections because he was afraid that the VR program would make the same mistakes.

[137] *Complètement perdu.*
[138] *Vieux réflexes.*

Another participant said that there is something "destabilizing"[139] about dictating a translation. He continued saying that, it is probably a reflex from someone who is not used to dictation, but that when he is dictating, he needs a few second break to think about what he is going to say. On this same topic, other participants affirm that ITD is more difficult because the sentence-building process is not the same. With this technique, they say, you have to stop and think about what you are going to say before dictating. One of them argued that it is a different working method, which is probably going to better suit certain personalities, because not everyone has the same habits, nor do they have the same way of thinking (hence, the interest of multimodal interaction is corroborated, since it provides various users with various interaction options).

One participant said that he felt "very insecure"[140] performing this task, that his "self-confidence was at its lowest."[141] He explained, for instance, that when he is typing a translation, and makes a mistake, "it's not a big deal,"[142] but that during the experiment making a mistake was "stressful".[143]

Another reason why the task is difficult is because the translator has to do all the information searches first to know in advance what he is going to say, and avoid "making the system go nuts"[144] because of hesitations, said one participant. Another

[139] *Déstabilisant.*
[140] *Très insecure.*
[141] *J'avais zéro confiance en moi.*
[142] *Ce n'est pas grave.*
[143] *Stressant.*
[144] *Faire capoter le système.*

participant said that it was “a bit tiresome”[145] having to stop dictating to search information. Likewise, another translator said that this task was difficult because there were many commands at the same time to integrate and to think about, including commands for turning the system on and off. Lastly, while acknowledging the difficulty of the task, one participant said that the more he had done performing the translation task, the easier it became. He said that at first he was afraid not to be able to translate using such an environment. He said that “it’s pretty complicated”[146] but less than he thought. “You get used to it pretty quickly.”[147]

The touch-screen interaction was a cause of difficulty for many participants. They say that they found it difficult to delete one character, or to add or remove a space, or to place the cursor within one word to insert one character for instance. Later, they acknowledged that this would not have been so difficult had they thought about zooming in the document with their fingers for the touch interaction to be more accurate. One participant concluded saying that it was “the old reflex of using the mouse”[148] that was difficult to let go in one hour.

In summary, most translators in our sample found the task to be difficult mostly because they were using this technique and these tools to translate for the very first time. However difficult the interaction was, there is a tendency to think that the more

[145] *Un peu tannant.*
[146] *C'est assez compliqué.*
[147] *On s’habitue rapidement.*
[148] *Le vieux réflexe de la souris.*

they will use this technique and these tools, the easier the task will become. As one participant put it: “I’m pretty much convinced that I could adapt to [ITD] very quickly and never want to type ever again.”[149] Now, let us see that, despite of the difficulties described above, there was a general sense of success among participants.

6.3.2.5. Sense of Success

In the interview, no particular question was asked to learn specifically about the participants’ sense of success. However, some of the statements may provide hints about the participants´ overall experience and satisfaction with the experimental environment.

One participant said that, despite all the limitations of the technology and the difficulties he had, the prototypical ITD environment was “impressive”[150], and that it was “pleasant to see”[151] that in one hour he managed to become acquainted with new techniques: “It is a method with a bright future ahead,”[152] he said. Another translator said that, in little time, the results were “very good”[153]. He continued explaining that he did not know the technology has progressed so much, that he still had his impressions from the past: that you needed a “significant amount of effort”[154] to train the tool. He concluded saying that if someone asks him: “would you want to work like this from now on?” he would say “yes”: “Now I get why people want to use it.”[155]

[149] *Je suis assez convaincu que je pourrais très rapidement m'habituer et ne vouloir taper plus jamais.*
[150] *Impressionnant.*
[151] *Agréable de voir.*
[152] *C'est une méthode qui a de l'avenir.*
[153] *Vraiment bien.*
[154] *Faire énormément d’efforts.*
[155] *Je comprends pourquoi les gens veulent l’utiliser.*

Another participant said he had been "extremely amazed"[156] by the tools, and that he was glad to have been able to take part in this study. He affirmed that he sees the potential with this type of environment, either to have it as a main tool or to have the choice according to the context. Indeed, he says, language professionals are those who need to choose the tools that suit them best.

Other indicators of the users' sense of success include comments such as "I find it really interesting";[157] "it's all about making adjustments"[158] because it was still "pretty impressive";[159] "In general, it went well for a first try";[160] "I was amazed to see to what extent the voice recognition software adapted to my voice; it made only a few mistakes";[161] and "the technology is there, and it gives pretty good results"[162].

Most of the comments above are in regard to voice input specifically. However, some participants reflected beyond VR and mentioned specifically the advantages of the multimodal configuration: "[multimodality] gives you choices",[163] said one participant, while another one said that "it is great that we are monitoring technological change [...]. We are living in a world that is based on touch interactions;

---

[156] *Extrêmement étonné.*
[157] *Je trouve ça vraiment intéressant.*
[158] *Tout est question d'ajustement.*
[159] *Assez impressionnant.*
[160] *En général, ça s'est bien passé pour un premier essai.*
[161] *J'ai été étonné de voir à quel point le logiciel de reconnaissance vocale s'est adapté; il y avait peu d'erreurs.*
[162] *La technologie est rendue là, elle donne de très bons résultats.*
[163] *Ça donne le choix.*

it is thus great that we have working tools that are compatible with that."[164] Another translator said that the greatest benefit he can see is that one can go find tools quickly (referring most certainly to the multimodal command combining touch-and-voice input). He said that it is "more fun"[165] to have an environment that is more interactive; either you're using voice, touch or the virtual keyboard when needed, rather than having just a keyboard and a mouse. "I think we only stand to gain from it,"[166] he concluded. Similarly, one of the participants affirmed that "a multi-device environment like this […], I think that it would truly be the total optimization for the translation industry".[167]

Although no generalizations can be formulated based on our sample, the multimodal ITD environment appears to provide translators with a better TX than the traditional KM environment, despite the occasional cases of lack of enthusiasm, the rare negative comments and the expressions of task difficulty by some participants. Thus, it can be said that my hypothesis was confirmed. Nonetheless, further experiments will be needed to build further knowledge about the TX with emerging input modes and computer devices.

[164] *C'est bien qu'on suive l'évolution des technologies […]. On est rendu dans un monde tactile, c'est bien qu'on ait des outils de travail qui sont compatibles avec ça.*

[165] *Plus le fun.*

[166] *Je pense qu'on a tout à gagner.*

[167] *Un environnement multi-appareil comme ça* […], *je pense que ça serait vraiment comme l'optimisation totale pour la traduction.*

6.3.3. Concluding Remarks on Qualitative Data Analysis

In this section, I presented subjective users' attitudes and satisfaction measures. While most of the objective measures presented in the previous section pointed towards a better quality-in-use of the ITD environment, it was necessary to complement such data with qualitative data from interviews with participants. By studying subjective measures, a bigger picture of the TX was observable. The qualitative data from the interviews did inform us about the participants' perceptions of and responses to the use of a KM workstation and the prototypical ITD environment, and the anticipated use of an optimized version of the ITD environment. In addition to quantitative data, qualitative data from real users, collected in real-life or simulated scenarios, can indeed help tool designers and developers in making informed decisions about various aspects of the UI of existing and new tools.

Chapter 7 — Conclusions and Future Work

"I haven't failed 10,000 times. I have successfully found 10,000 ways that won't work." — Thomas A. Edison

"When you have exhausted all possibilities, remember this: you haven't." — Thomas A. Edison

Translation is perhaps one of the oldest practices in human history. For millennia, humans have communicated with peoples from other tribes, cultures, countries, continents, languages and times; and translators have been essential to mediating such communications.

With the emergence of alphabets, writing systems and writing tools has come the emergence of written translation as well. Over the centuries, translators have adopted different tools according to ongoing developments in science and technology: from stone-engraving tools and ink pens, to typewriters and, more recently, personal computers.

Computers have served professional translators for over three decades now, not only by allowing them to type translations but also by enabling them to access information and to exploit a number of technological aids and applications, developed to make translators' work easier and/or more efficient. However, since the turn of the new millennium, the translation industry seems to be surprisingly lagging behind the evolution of writing and information tools, now being designed more and more with end users in mind. Only a handful of research efforts have consulted with translators when designing and implementing tools; most efforts have focused mainly on improving the efficiency and performance of tools, with little or no consideration for humans' changing needs and preferences in terms of accessing, using and communicating information.

A definite new turn in translation technology is needed, with the human translator as the central axis of investigation. Shneiderman (2015, p. ix) affirms that:

> [t]echnology designers who shape user experiences seek to smooth the path for novices and serve the demanding needs of experts. This was true for fifteenth-century book designers, nineteenth-century train designers, and twenty-first-century smartphone designers. Their innovative designs emerged from a deep empathy for people, sensitivity to diverse social contexts, and imaginative sparks to create new ways of thinking about technology.

This doctoral study creates new ways of thinking about translation technology. It provides a basis for well-grounded research on translator-computer and translator-information interaction and, more specifically, for the design and development of an interactive translation dictation environment. Such an interactive system is expected to support professional translators' cognitive functions, performance and well-being.

The interaction of translators with multimodal interfaces will make the translation task more human. In the words of Oviatt (2012, p. 423):

> "[m]ultimodal interfaces are just beginning to model human-like sensory perception and communication patterns. They are recognizing and identifying actions, language and people […]. They literally reflect and acknowledge the existence of human users, empower them in new ways, and create for them a "voice".

This thesis followed the following path: First, I presented the evolution of the interaction between translators, tools, computers and information, and the evolution of reflections on this matter among the research community. Then, I discussed the advantages and challenges of current translation techniques, including typed translation and various forms of translation dictation. Subsequently, I explored the literature on bilingual voice recognition technology and multimodal interaction, and provided a discussion on these technologies from a translation point of view. In light of this, I described interactive translation dictation in more detail. For the experimental portion of this study, I discussed the challenges of measuring usability and presented the measures used here. Next, I looked at some data-collection techniques used in translator-centred research, and then provided a detailed description of the methodology used to collect empirical quantitative and qualitative data. Lastly, I presented and analyzed both objective usability data using measures such as expert assessment of the outcome, task times and usability patterns; and subjective data using measures such as users' perception of the outcome, subjectively experienced time and perception of task difficulty. I also attempted to demonstrate how the inclusion of qualitative data in usability studies, and a combined analysis of these and quantitative data, may provide a richer picture of the actual interaction experience than quantitative data analysis alone.

As explained in the introductory chapter, research projects in human-computer interaction seek to design or redesign a particular computing technology in order to improve upon or enhance a given experience or to create a different experience than before (Harper et al. 2008, 58). Based on the findings gathered from initial user studies, researchers begin to think about how to design something better (*ibid*.). The specific goal of my study was to investigate the interaction of translators with voice-and-touch multimodal interfaces to provide a basis for designing, developing and testing the working prototype of an interactive translation dictation environment.

This doctoral thesis should contribute to the advancement of knowledge mainly in the fields of translation studies, multimodal interaction, human-computer interaction and human-information interaction in general. One of the purposes was to demonstrate that any research on tools and technology needs to be more inclusive of all the various measures and aspects of usability, and more considerate of the user experience and the changing realities of the interactions between humans, computers and information. New realities lead to new tools and new contexts of use (e.g. home technology, ubiquitous computing, cloud computing, on-line learning), and these new contexts require new measures, as well as new data collection and analysis approaches to adequately capture what is considered important in particular contexts and for particular users. Schneiderman (2015, p. ix) argues that "[e]ach generation of designers faces fresh opportunities to remake human experiences in ways that will be easier, safer, more enjoyable, and even more compelling than the past." The next challenge in translator-computer and translator-information interaction is to design for different interaction channels (that is, different types of devices and user interfaces), different users, and different contexts of use. To effectively achieve this, the conceptual design path must be followed (Parush, 2015).

Together with my work at the master's level, this study constitutes an initial literature review and user research, which is only the first stage of the long process of envisioning, designing, developing, testing and commercializing an interactive system or product. The present study was carried out with the following research question in mind: Does the physical prototype of an interactive translation dictation environment consistently provide translators with a better translator experience than a traditional keyboard-and-mouse environment, when preparing, producing and revising a translation? The combined analysis of objective and subjective usability measures collected *does* suggest that the translator experience is better with the experimental environment than with the traditional workstation, considering the translation process as a whole. In other words, our sample indicates that translators' perceptions and responses are more positive, generally speaking, in interactive translation dictation. I am confident that the objectives of the study were successfully met, and that the hypothesis was confirmed. Now, in order for interactive translation dictation to be fully integrated to the translation profession, much work must be done.

First of all, future work will need to cover various elements that were left out of the analyses in this thesis. Given the financial, ethical and time constraints of my doctorate, only a portion of the quantitative and qualitative data collected could be presented and analyzed here. My personal work in the near future will involve examining these data in further detail, including taking a closer look at video and screen recordings as well as at interview answers. By further examining these sources of data, I should be able to determine what the particularities of the interaction with the experimental environment were, and where exactly the speed bumps and frustrations in this interaction lay. Additionally, I should be able to determine what tool- and device-specific difficulties are involved in carrying out translation tasks, and what—according to potential users of an interactive translation dictation environment—external difficulties prevent such an environment and technique from being fully integrated into the profession. Lastly, a closer look at the interview data can also provide me with clues to what users want (i.e. what features of an interactive translation dictation environment are "nice to have"), which of the tested applications and devices have a "wow" factor, and which ones should be ruled out from the design of an interactive translation dictation environment.

Beyond my own work, translation technology scholars and researchers will need to engage in conceptual design, by building conceptual models. What does a researcher need to consider before designing an interactive system in order to ensure that the system will provide the user with a positive experience? (Parush, 2015, p. xi). By engaging in conceptual design, researchers can make an interactive system more usable, saving various stakeholders significant time, money and other resources as a result (*ibid*., xiii). Hence, researchers "avoid being part of the proliferation of unusable systems and suffer the costs of redesigning those systems to make them effective" (*ibid*.). Parush (*ibid*., pp. 43–49) explains that the discussion surrounding the importance of engaging in conceptual design focuses on the following five fundamental human psychology and performance factors, which play a role in any interaction:

1. Mental models ("If you understand it, you can use it!")

2. Location awareness ("If you know where you are, you can reach your destination!")

3. Visual search effectiveness ("If you can find what you look for, you can accomplish your goal!")

4. Operational load ("If you do fewer actions, your effort is reduced!")

5. Working memory load ("If you are not required to remember much, you can do more!").

These factors can have a great impact on the overall user experience and on the quality-in-use of the interface being designed, tested and implemented.

Moreover, future empirical research in translation studies, and specifically on interactive translation dictation, will need to investigate in more detail particular interaction modes such as bilingual voice recognition, on the one hand, and stylus or digital pen input, on the other hand. New user studies could explore in detail which tasks, in a regular translation workflow, are done in one language or the other, and how those tasks could be automatized or optimized by enabling bilingual voice interaction. In addition, to complement an experiment such as the pilot study reported in appendix C, a second-phase experiment could be carried out with a touch-enabled stylus interface, in which participants simulate the use of a pen to prepare a source text printed on paper before dictating a translation.

In addition, translation studies researchers will need to investigate more boldly the changing nature of computer devices, as we move towards a new computing era.

Mobile and cloud computing are changing the way humans interact with computers and information. My experiment was performed using a reclining all-in-one desktop computer and a tablet, both devices connected to a wireless Internet network. Apart from the computer's power adaptor, no wires were involved in the experimental setup. No headset was needed for the voice recognition on any of the devices. In addition, even though the desktop computer's voice recognition system was installed on the device, the technology is already moving into the cloud, and its accuracy can be comparable to that of embedded systems (Zapata & Kirkedal, 2015). Thus, a totally wireless-and-mobile experiment similar to the one presented in this thesis is already possible, for example, with a tablet and a smartphone only.

Future investigations will also need to consider the integration of translation memories and machine translation technology to an interactive translation dictation environment. Translation memory systems and machine translation were excluded from the experimental portion of this thesis. While possible, an experiment that combines these applications and emerging input modes is rather challenging and beyond the scope of this work. In this study, I explored emerging input modes such as voice, touch and stylus to produce from-scratch translations only. In an interactive translation dictation environment, however, translation memories and machine translation may indeed augment the workflow, and improve the efficiency of the interactive translation dictation environment in the background, by using a hybrid approach similar to those explored in the 1990s and in the first decade of the present century.

Last but not least, translation studies scholars will need to explore the pedagogical potential and challenges of interactive translation dictation. One of the challenges identified in my previous research (Zapata Rojas, 2012) and in the present study[168] is the lack of formal training in oral translation techniques and in the use of emerging technologies. Universities play an increasingly crucial role in efficiently integrating new technologies such as voice recognition and mobile devices into the translator's toolbox (Mees et al., 2013, p. 152). And, in this integration, translators will need to learn to dictate efficiently. They will need to adopt "completely new translation techniques" (Gouadec, 2007, p. 286). I have personally been working on developing and testing an interactive translation dictation introductory workshop, which I have taught in various universities in Canada and abroad. Furthermore, I co-authored a paper on the need to integrate interactive translation dictation into translator training programs.[169]

The road to interactive translation dictation integration is long yet steady. One thing is certain: in all probability, interactive translation dictation *will* be the norm in translation teaching and practice. There is indeed a bright future for this technique, for multimodal interfaces, and for mobile and cloud computing in the translation field. Dee Hock, as quoted at the very beginning of this thesis, tells us that "[t]he problem

[168] Experimental data informing on the pedagogical challenges of interactive translation dictation were excluded, since they go beyond the scope of the study. The data were presented at a panel on the teaching of computer-aided translation tools within the framework of the 5th IATIS conference, held in Belo Horizonte, Brazil, in July 2015.

[169] As of this writing, the paper is under peer review for publication in a prestigious international translation journal.

is never how to get new, innovative thoughts into your mind, but to get old ones out." Constant change and innovation in translation technology is a must.

No one denies that translators—and humans in general—are increasingly dependent on computer tools. Thus, in this age of translation technology, mobile and cloud computing and ubiquitous information, research on speech technologies and multimodal interaction will become increasingly important within translation studies. Experiments that explore voice, touch and pen input (and even other emerging interaction modes such as gaze, gesture and brain input) will play a crucial role in the design and development of new tools that are user-friendly and adapted to translators' needs and to the changing reality of the translation industry, reflecting the continuous transformations in the way humans and machines interrelate.

Bibliography

Alabau, V., & Casacuberta, F. (2012). Study of Electronic Pen Commands for Interactive-Predictive Machine Translation. In *Proceedings of the International Workshop on Expertise in Translation and Post-editing Research and Applications*. Retrieved from https://prhlt.iti.upv.es/aigaion2/attachments/epen-gestures.pdf-7d6caac5dec090b6148c203b2305cbf1.pdf

Alabau, V., & Leiva, L. (2014). Proofreading Human Translations with an E-pen. In *Workshop on Humans and Computer-assisted Translation* (pp. 10–15). Association for Computational Linguistics.

Alabau, V., & Martínez, C. D. (2006). Bilingual Speech Recognition in Two Phonetically Similar Languages. In *Proceedings of the VI Jornadas en Tecnologías del Habla*. Retrieved from http://lorien.die.upm.es/~lapiz/rtth/JORNADAS/IV/finals/4jth_118.pdf

Alabau, V., Ortiz-Martínez, D., Sanchis, A., & Casacuberta, F. (2010). Multimodal interactive machine translation. In *International Conference on Multimodal Interfaces and the Workshop on Machine Learning for Multimodal Interaction on - ICMI-MLMI '10*. New York, New York, USA: ACM Press. doi:10.1145/1891903.1891960

Alabau, V., Rodríguez-Ruiz, L., Sanchis, A., Casacuberta, P., & Martínez-Gómez, F. (2011). On Multimodal Interactive Machine Translation Using Speech Recognition Categories and Subject Descriptors. In *Proceedings of the 13th International Conference on Multimodal Interfaces* (pp. 129–136).

Alabau, V., Sanchis, A., & Casacuberta, F. (2011). Improving On-line Handwritten Recognition using Translation Models in Multimodal Interactive Machine Translation. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies* (Vol. 47, pp. 389–394).

Arnold, D. (2003). Why Translation Is Difficult for Computers. In H. Somers (Ed.), *Computers and Translation: A Translator's Guide* (pp. 119–142). Amsterdam: John Benjamins.

Aubert, B., Cohendet, P., & Montreuil, B. (2010). *L'innovation et les technologies de l'information et des communications*. Québec: CEFRIO.

Bangalore, S., & Johnston, M. (2009). Robust Understanding in Multimodal Interfaces. *Computational Linguistics*, *35*(3), 345–397.

Barabasi, L. (2012). Une théorie de la complexité s'appuiera sur la science des réseaux. *La Recherche*, *464*(May), 38–41.

Barabé, D. (2013). Société, technologie et traduction : perspectives et impacts. *The Journal of Specialised Translation*, *19*, 41–31.

Bargas-Avila, J. A., & Hornbæk, K. (2011). Old Wine in New Bottles or Novel Challenges ? A Critical Analysis of Empirical Studies of User Experience. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI 2011)* (pp. 2689–2698). New York, New York, USA: ACM Press.

Barrachina, S., Bender, O., Casacuberta, F., Civera, J., Cubel, E., Khadivi, S., … Vilar, J.-M. (2009). Statistical approaches to computer-assisted translation. *Computational Linguistics*, *35*(1), 3–28. Retrieved from http://eprints.pascal-network.org/archive/00005438/

Benis, M. (2002). Softly spoken or hard of hearing? *Language International*, *14*(3), 26–29.

Bevan, N. (1995). Measuring Usability as Quality of Use. *Software Quality Journal*, *4*, 115–150.

Bolt, R. A. (1980). "Put-that-there": Voice and gesture at the graphics interface. In *Proceedings of the SIGGRAPH'80* (pp. 262–270). New York: ACM Press.

Borja, A. (2008). Corpora for Translators in Spain. The CDJ-GITRAD corpus and the GENTT Project. In M. Rogers & G. Anderman (Eds.), *Incorporating Corpora: The Linguist and the Translator* (pp. 243–265). Clevedon: Multilingual Matters.

Bowker, L. (2002). *Computer-Aided Translation Technology: A Practical Introduction*. Ottawa: University of Ottawa Press.

Bowker, L. (2006). Translation Memory and "Text." In L. Bowker (Ed.), *Lexicography, Terminology and Translation: Text-based Studies in Honour of Ingrid Meyer* (pp. 175–187). Ottawa: Les Presses de l'Université d'Ottawa.

Bowker, L. (2011). Off the Record and on the Fly: Examining the Impact of Corpora on Terminographic Practice in the Context of Translation. In *Lexicography, Terminology and Translation: Text-based Studies in Honour of Ingrid Meyer* (pp. 175–187). Ottawa: University of Ottawa Press. Retrieved from http://site.ebrary.com/lib/oculottawa/docDetail.action?docID=10139181

Bowker, L., & Buitrago Ciro, J. (2015). User eXperience and Translatability Viewed through the Lens of a Triple Constraint: Time, Cost and Quality. In *Proceedings of the Annual Conference of CAIS*. Ottawa, June 3-5, 2015.

Bowker, L., McBride, C., & Marshman, E. (2008). Getting more than you paid for? Considerations in integrating free technologies into translator training programs. *Redit*, *1*(1), 26–47. Retrieved from http://www.redit.uma.es/Archiv/v1_2008/Bowker_et_al.pdf

Brousseau, J., Drouin, C., Foster, G., Isabelle, P., Kuhn, R., Normandin, Y., & Plamondon, P. (1995). French speech recognition in an automatic dictation system for translators: The TransTalk project. In *Proceedings of Eurospeech'95*. Retrieved from http://www.iro.umontreal.ca/~foster/papers/ttalk-eurospeech95.pdf

Brunette, L., & Gagnon, C. (2013). Enseigner la révision à l'ère des wikis : là où l'on trouve la technologie alors qu'on ne l'attendait plus. *Journal of Specialized Translation*, *19*(January), 96–121. Retrieved from http://www.jostrans.org/issue19/art_brunette_gagnon.php

Budge, E. A. W. (1989). *The Rosetta Stone*. New York: Dover.

Callon, M. (2006). Sociologie de l'acteur réseau. In A. Akrich, M. Callon, & B. Latour (Eds.), *Sociologie de la traduction* (pp. 267–276). Paris: Presses de l'École des Mines.

Carl, M. (2012a). Translog - II: a Program for Recording User Activity Data for Empirical Reading and Writing Research. In *Proceedings of the Eight International Conference on Language Resources and Evaluation* (pp. 4108–4112).

Carl, M. (2012b). Translog - II: a Program for Recording User Activity Data for Empirical Reading and Writing Research. In *Proceedings of the Eight International Conference on Language Resources and Evaluation* (pp. 4108–4112). Istanbul, May 21-27, 2012: European Language Resources Association.

Carl, M., Bangalore, S., & Schaeffer, M. (Eds.). (2016). *New Directions in Empirical Translation Process Research*. New York: Springer.

Carl, M., Dragsted, B., & Jakobsen, A. L. (2011). On the Systematicity of Human Translation Processes. In *Actes de la conférence de Tralogy, Paris, le 3 mars 2011*. Retrieved from http://lodel.irevues.inist.fr/tralogy/index.php?id=103

Chatterjee, I., Xiao, R., & Harrison, C. (2015). Gaze + Gesture : Expressive, Precise and

Targeted. In *Proceedings of the 17th International Conference on Multimodal Interaction ICMI15* (pp. 131–138).

Church, K. W., & Hovy, E. H. (1993). Good Applications for Crummy Machine Translation. *Machine Translation*, *8*, 239–258. Retrieved from http://journals1.scholarsportal.info/tmp/15390858948988344755.pdf

Ciobanu, D. (2014). Of Dragons and Speech Recognition Wizards and Apprentices. *Revista Tradumàtica*, (12), 524–538.

Cohen, P., Johnston, M., McGee, D., Oviatt, S. L., Clow, J., & Smith, I. (1998). The Efficiency of Multimodal Interaction: A Case Study. In *Proceedings of the ICSLP'98*. Retrieved from http://web.media.mit.edu/~earroyo/thesis-backup/Thesis-prep/multimodal-interfaces/Alredy-printed/the-efficiency-of-multimodal-interaction.pdf

Cohen, P. R., Moran, D. B., Ave, R., Park, M., Sullivan, J. W., & Schlossberg, J. L. (1989). Synergistic Use of Direct Manipulation and Natural Language. In *CHI'89 Proceedings* (pp. 227–233). Retrieved from http://dl.acm.org/citation.cfm?id=67494

Coleman, M. L. (1969). Text editing on a graphic display device using hand-drawn proofreader's symbols. In *Pertinent Concepts in Computer Graphics, Proc. 2nd Univ. Illinois Conf. on Computer Graphics* (pp. 283–290).

Daiute, C. (1985). *Writing and Computers*. Reading, Massachusetts: Addison-Wesley.

De Schaetzen, C. (1995). Bureautique et télé-bureautique en traduction. *Meta*, *40*(4), 682–693. Retrieved from http://www.erudit.org/revue/meta/1995/v40/n4/004040ar.pdf

Désilets, A., Melançon, C., Patenaude, G., & Brunette, L. (2009). How Translators Use Tools and Resources to Resolve Translation Problems: an Ethnographic Study. *Beyond Translation Memories: New Tools for Translators, MT-Summit Workshop, Ottawa, Le 29 Août 2009*, *9*, 1–8. Retrieved from http://www.mt-archive.info/MTS-2009-Desilets-2.pdf

Désilets, A., Stojanovic, M., Lapointe, J.-F., Rose, R., & Reddy, A. (2008). Evaluating Productivity Gains of Hybrid ASR-MT Systems for Translation Dictation. Retrieved from http://www.mt-archive.info/IWSLT-2008-Desilets.pdf

Dieguez-Tirado, J., García-Mateo, C., Docio-Fernández, L., & Cardenal-López, A. (2005). Adaptation Strategies for the Acoustic and Language Models in Bilingual Speech Transcription. In *Proceedings of the ICSASSOP´05* (pp. 833–

836).

Dourish, P. (2001). *Where The Action Is: The Foundations of Embodied Interaction* (1st ed.). Cambridge, Massachusetts: The MIT Press.

Dragsted, B., & Hansen, G. (2009). Exploring Translation and Interpreting Hybrids. The Case of Sight Translation. *Meta*, *54*(3), 588–604. Retrieved from http://www.erudit.org/revue/meta/2009/v54/n3/038317ar.pdf

Dragsted, B., Hansen, I. G., & Selsøe Sørensen, H. (2009). Experts Exposed. *Copenhaguen Studies in Language*, *38*, 293–317.

Dragsted, B., Mees, I. M., & Hansen, I. G. (2011). Speaking your translation: students' first encounter with speech recognition technology. *Translation & Interpreting*, *3*(1), 10–43. Retrieved from http://www.trans-int.org/index.php/transint/article/viewFile/115/87

Duval, C. (2012). L'impact du Web en 4 questions. *La Recherche*, *467*(September), 46–50.

Dymetman, M., Brousseau, J., Foster, G., Isabelle, P., Normandin, Y., & Plamondon, P. (1994). Towards an Automatic Dictation System for Translators: the TransTalk Project. In *Fourth European Conference on Speech Communication and Technology* (p. 4). Retrieved from http://arxiv.org/abs/cmp-lg/9409012

Ehrensberger-Dow, M., & Massey, G. (2014). Internal and external resources: constraints on creativity? In *Writing Research Across Borders III. Writing, Language, and New Media: Challenges and Possibilities Across Educational Contexts* (pp. 19–22). Paris: WRAB 2014.

Ehrensberger-Dow, M., & O'Brien, S. (2015). Ergonomics of the Translation Workplace: Potential for Cognitive Friction. *Translation Spaces*, *4*(1), 98–118. doi:10.1075/ts.4.1.05ehr

Enríquez Raído, V. (2014). *Translation and Web Searching*. New York: Routledge.

Ericsson, K. A., & Simon, H. A. (1984). *Protocol Analysis: Verbal Reports as Data*. Cambridge, Massachusetts: MIT Press.

Fidel, R. (2012). *Human Information Interaction. An Ecological Approach to Information Behaviour*. Cambridge, Massachusetts: MIT Press.

Foster, G., Isabelle, P., & Plamondon, P. (1998). Target-text mediated interactive machine translation. *Machine Translation*, *12*, 175–194.

Garcia, I. (2009). Beyond Translation Memory: Computers and the Professional

Translator. *Journal of Specialised Translation*, *12*, 199–214. Retrieved from http://www.jostrans.org/issue12/art_garcia.pdf

Garcia-Martinez, M., Singla, K., Tammewar, A., Mesa-Lao, B., Thakur, A., Anusuya, M. A., … Bangalore, S. (2014). SEECAT: ASR & Eye-tracking Enabled Computer-Assisted Translation. In *Proceedings of the 17th Annual Conference of the European Association for Machine Translation* (pp. 81–88).

Gershon, N. (1995). Human Information Interaction. In *4th International World Wide Web Conference*.

Gingold, K. (1978). The Use of Dictation Equipment in Translation. In P. A. Horguelin (Ed.), *La traduction, une profession. Actes du VIIIe congrès mondial de la Fédération internationale des traducteurs* (pp. 444–448). Ottawa: Conseil des traducteurs et interprètes du Canada.

Girol, D., Lopez-Cozar, R., Callejas, Z., Espejo, G., & Ábalos, N. (2012). On the Development of Adaptive and User-Centered Interactive Multimodal Interfaces. In U. S. Tiwary & T. J. Siddiqui (Eds.), *Speech, Image and Language Processing for Human Computer Interaction: Multi-Modal Advancements* (pp. 262–291). Hershey, Pennsylvania: Information Science Reference.

Gouadec, D. (2002). *Profession : traducteur*. Paris: La Maison du dictionnaire.

Gouadec, D. (2007). *Translation as a profession*. Amsterdam: John Benjamins.

Gouanvic, J.-M. (1976). Esquisse d'une méthode de traduction au dictaphone. *Meta*, *21*(4), 252–255. Retrieved from http://www.erudit.org/revue/meta/1976/v21/n4/002951ar.pdf

Harper, R., Rodden, T., Rogers, Y., & Sellen, A. (2008). *Being Human: Human-Computer Interaction in the Year 2020*. Cambridge.

Hartmann, B., Klemmer, S. R., Bernstein, M., Abdulla, L., Burr, B., Robinson-Mosher, A., & Gee, J. (2006). Reflective physical prototyping through integrated design, test, and analysis. *Proceedings of the 19th Annual ACM Symposium on User Interface Software and Technology - UIST '06*, 299. doi:10.1145/1166253.1166300

Haton, J.-P., Cerisara, C., Fohr, D., Laprie, Y., & Smaili, K. (2006). *Reconnaissance automatique de la parole. Du signal à son interprétation.* Paris: Dunod (UniverSciences).

Havens, T., & Lotz, A. D. (2012). *Understanding Media Industries*. New York: Oxford

University Press.

Hayes, P. (1997). What is a Computer? An Electronic Discussion. *The Monist*, *80*(3), 389–404.

Hétu, M.-P. (2012). Le travail au dictaphone, une solution ergonomique? *Circuit*, *116*(été 2012), 23.

Hewett, T. T., Baecker, R., Card, S., Carey, T., Gasen, J., Mantei, M., … Verplank, W. (1992). *ACM SIGCHI Curricula for Human-Computer Interaction*. New York: The Association for Computing Machinery. Retrieved from http://www2.parc.com/istl/groups/uir/publications/items/UIR-1992-11-ACM.pdf

Hornbæk, K. (2006). Current practice in measuring usability: Challenges to usability studies and research. *International Journal of Human-Computer Studies*, *64*(2), 79–102. doi:10.1016/j.ijhcs.2005.06.002

Hornecker, E., & Buur, J. (2006). Getting a Grip on Tangible Interaction: A Framework on Physical Space and Social Interaction. In *CHI'06 Proceedings* (pp. 437–446). Montreal. doi:http://doi.acm.org/10.1145/1124772.1124838

Inkpen, K. M. (2001). Drag-and-drop versus point-and-click mouse interaction styles for children. *ACM Transactions on Computer-Human Interaction*, *8*(1), 1–33.

ISO, 9241-11:1998. (1998). *Ergonomic requirements for office work with visual display terminals (VDTs) -- Part 11: Guidance on usability*. International Organization for Standarization (ISO), Switzerland.

ISO, 9241-210:2010. (2010). *Ergonomics of human system interaction - Part 210: Human-centred design for interactive systems*. International Organization for Standarization (ISO), Switzerland.

Jakobsen, A. L. (2003). Effects of think aloud on translation speed, revision and segmentation. In F. Alves (Ed.), *Triangulating Translation. Perspectives in Process Oriented Research* (pp. 69–95). Amsterdam: Benjamins.

Jakobsen, A. L. (2011). Tracking translators' keystrokes and eye movements with Translog. In C. Alvstad, A. Hild, & E. Tiselius (Eds.), *Methods and Strategies of Process research. Integrative approaches in Translation Studies* (pp. 37–55). Amsterdam: Benjamins.

Jiménez Ivars, A. (1998). *La traducción a vista. Un análisis descriptivo*. Universitat Jaume I. Retrieved from http://www.tesisenxarxa.net/TESIS_UJI/AVAILABLE/TDX-0519103-121513/jimenez-tdx.pdf

Jiménez Ivars, A., & Hurtado Albir, A. (2003). Variedades de traducción a la vista. Definición y clasificación. *Trans Revista de Traductología*, *7*, 47–57. Retrieved from http://www.trans.uma.es/trans_07.html

Jurafsky, D., & Martin, J. H. (2009). *Speech and language processing: An introduction to natural language processing, computational linguistics, and speech recognition* (2nd ed.). Upper Saddle River, New Jersey: Pearson Prentice Hall.

Karamanis, N., Luz, S., & Doherty, G. (2011). Translation practice in the workplace: A contextual analysis and implications for machine translation. *Machine Translation*, *25*(1), 35–52.

Koehn, P., & Germann, U. (2014). The Impact of Machine Translation Quality on Human Post-editing. In *Workshop on Humans and Computer-assisted Translation* (pp. 38–46).

Koehn, P., & Haddow, B. (2009). Interactive assistance to human translators using statistical machine translation methods. In *Proceedings of MT Summit XII* (pp. 73–80). Ottawa.

Kumar, S. C., & Say Wey, F. (2003). A bilingual speech recognition system for English and Tamil. In *Proceedings of the ICICSP´03* (pp. 1641–1644).

Kurzweil, R. (2013). *How to Create a Mind. The Secret of Human Thought Reveailed*. New York: Penguin.

Lambert, S. (1989). La formation d'interprètes : la méthode cognitive. *Meta*, *34*(4), 736–744. Retrieved from http://www.erudit.org/revue/meta/1989/v34/n4/003379ar.pdf

Lambert, S. (1991). Aptitude Testing for Simultaneous Interpretation at the University of Ottawa. *Meta*, *36*(4), 586–594. Retrieved from http://www.erudit.org/revue/meta/1991/v36/n4/003383ar.pdf

Lambert, S. (2004). Shared attention during sight translation, sight interpretation and simultaneous interpretation. *Meta*, *49*(2), 294–306. Retrieved from http://www.erudit.org/revue/meta/2004/v49/n2/009352ar.pdf

Langlais, P., Sauvé, S., Foster, G., Macklovitch, E., & Lapalme, G. (2000). Evaluation of TransType, a computer-aided translation typing system: a comparison of a theoretical-and a user-oriented evaluation procedures. In *Conference on*

Language Resources and Evaluation (LREC) (p. 8). Retrieved from http://www.iro.umontreal.ca/~foster/papers/tteval-lrec00.pdf

Laroque-Divirgilio, L. (1980). Sondage sur les services de traduction dans l'entreprise. *Meta*, *25*(2).

Laroque-Divirgilio, L. (1981). La traduction au magnétophone. *Meta*, *26*(4), 398–403. Retrieved from http://www.erudit.org/revue/meta/1981/v26/n4/002573ar.pdf

Latour, B. (1996). Portrait de Gaston Lagaffe en philosophe des techniques. In *Portrait de Gaston Lagaffe en philosophe des techniques. Petites Leçons de sociologie des sciences.* (pp. 15–24). Paris: La Découverte.

Latour, B. (1997). On actor-network theory. A few clarifications. Retrieved from http://www.nettime.org/Lists-Archives/nettime-l-9801/msg00019.html

LeBlanc, M. (2013). Translators on translation memory (TM). Results of an ethnographic study in three translation services and agencies. *The International Journal for Translation & Interpreting Research*, *5*(2), 1–13.

Lee, B., Isenberg, P., Riche, N. H., & Carpendale, S. (2012). Beyond Mouse and Keyboard: Expanding Design Considerations for Information Visualization Interactions. *IEEE Transactions on Visualization and Computer Graphics*, *18*(12), 2689–2698.

Lei, X., Senior, A., Gruenstein, A., & Sorensen, J. (2013). Accurate and Compact Large Vocabulary Speech Recognition on Mobile Devices. *Interspeech*, (August), 662–665. Retrieved from http://research.google.com/pubs/archive/41176.pdf

Leijten, M., & Van Waes, L. (2005). Writing with speech recognition: The adaptation process of professional writers with and without dictating experience. *Interacting with Computers; HCI and the Older Population*, *17*(6), 736–772.

Leijten, M., & Van Waes, L. (2013). Keystroke Logging in Writing Research: Using Inputlog to Analyze and Visualize Writing Processes. *Written Communication*, *30*(3), 358–392. Retrieved from http://wcx.sagepub.com/cgi/doi/10.1177/0741088313491692

Leiva, L., Alabau, V., & Vidal, E. (2013). Error-proof, high performance, and context gestures for interactive text edition. In *Proc. CHI EA* (pp. 1227–1232).

Llisterri, J. (2009a). Las tecnologías del habla en las lenguas románicas ibéricas. *Studies in Hispanic and Lusophone Linguistics*, *2*(1), 133–180.

Llisterri, J. (2009b). Les technologies de la parla. *Llengua, Societat I Comunicació. Revista de Sociolingüística de La Universitat de Barcelona*, *7*, 11–19.

Macklovitch, E. (2006). TransType2: The last word. In *Proceedings of the 5th International Con- ference on Languages Resources and Evaluation* (pp. 167–172). Retrieved from http://www.mt-archive.info/LREC-2006-Macklovitch.pdf

Macklovitch, E., & Russell, G. (2000). What's Been Forgotten in Translation Memory? In J. S. White (Ed.), *Envisioning Machine Translation in the Information Future* (pp. 137–146). Berlin: Springer Verlag.

Marchionini, G. (2008). Human-information interaction research and development. *Library & Information Science Research*, *30*(3), 165–174. Retrieved from http://www.ils.unc.edu/~march/Marchionini_Inf_interact_LISR_2008.pdf

Mariani, J. (2009). *Language and speech processing*. Hoboken: John Wiley & Sons.

Mariño, J. B., Pedrell, J., Moreno, A., & Nadeu, C. (2000). Monolingual and Bilingual Spanish-Catalan Speech Recognizers Developed from SpeechDat Databases. In *Proceedings of the International Worksop on Very Large Telephone Speech Databases,* (pp. 57–61).

Mees, I. M., Dragsted, B., Hansen, I. G., & Jakobsen, A. L. (2013). Sound effects in translation. *Target*, *25*(1), 140–154. Retrieved from http://openurl.ingenta.com/content/xref?genre=article&issn=0924-1884&volume=25&issue=1&spage=140

Mell, P., & Grance, T. (2011). The NIST Definition of Cloud Computing. National Institute of Standards and Technology Special Publication 800-145. Retrieved from http://csrc.nist.gov/publications/nistpubs/800-145/SP800-145.pdf

Mesa-Lao, B. (2014). Speech-Enabled Computer-Aided Translation: A Satisfaction Survey with Post-Editor Trainees. In *Workshop on Humans and Computer-assisted Translation* (pp. 99–103).

Misu, T. (2015). Visual Saliency and Crowdsourcing-based Priors for an In-car Situated Dialog System. In *Proceedings of the 17th International Conference on Multimodal Interaction ICMI15* (pp. 75–82).

Muegge, U. (2013). Teaching computer-assisted translation in the 21st century. In A.-K. D. Ende, S. Herold, & A. Weilandt (Eds.), *Alles hängt mit allem zusammen: Translatologische Interdependenzen. Festschrift für Peter A. Schmitt* (pp. 137–146). Berlin: Frank & Timme.

Navarro, G. (2001). A guided tour to approximate string matching. *ACM Computing Surveys*, *33*(1), 31–88.

Neal, J. G., & Shapiro, S. C. (1991). Intelligent multimedia interface technology. In *Intelligent User Interfaces* (pp. 11–43). New York: ACM Press.

Nielsen, J. (2012). Usability 101: Introduction to Usability. Retrieved from https://www.nngroup.com/articles/usability-101-introduction-to-usability/

O'Brien, S. (2009). Eye tracking in translation-process research: methodological challenges and solutions. *Copenhaguen Studies in Language*, *38*, 251–266.

O'Brien, S. (2012). Translation as human–computer interaction. *Translation Spaces*, *1*(1), 101–122. doi:10.1075/ts.1.05obr

O'Brien, S., & Ehrensberger-Dow, M. (2013, October). Ergonomic Issues at the Professional Translation Workplace. *Irish Translators' and Interpreters' Association Bulletin*, 5–6. Retrieved from https://www.zhaw.ch/storage/linguistik/forschung/uebersetzungswissenschaft/press-ergonomic-issues-professional-translation-workplace-2013.pdf

O'Brien, S., Moorkens, J., & Vreeke, J. (2014). Kanjingo – A Mobile App for Post-Editing. In *The 17th annual conference of the European Association for Machine Translation*.

O'Grady, W. (2004). Language: a preview. In W. O'Grady & J. Archibald (Eds.), *An Introduction to Contemporary Linguistic Analysis* (5th ed., pp. 1–11). Toronto: Pearson Education Canada.

Oviatt, S. (1999). Ten myths of multimodal interaction. *Communications of the Association for Computing Machinery*, *43*(9), 45–51.

Oviatt, S. (2012). Multimodal Interfaces. In J. A. Jacko (Ed.), *The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications* (3rd ed., pp. 415–429). New York: Lawrence Erlbaum Associates.

Oviatt, S. (2013). Interfaces for Thinkers : Computer Input Capabilities that Support Inferential Reasoning. In *Proceedings of the 15th ACM on International conference on multimodal interaction* (pp. 221–228).

Oviatt, S., & Cohen, P. R. (2000). Multimodal systems that process what comes naturally. *Communications of the ACM*, *43*(3), 45–53.

Oviatt, S., Cohen, P. R., Fong, M., & Frank, M. (1992). A rapid semi-automatic simulation technique for investigating interactive speech and handwriting. In

Proceedings of thw ICSLP'92 (pp. 1351–1354).

Oviatt, S., Cohen, P., & Wang, M. (1994). Toward interface design for human language technology: Modality and structure as determinants of linguistic complexity. *Speech Communication*, *15*, 283–300.

Oviatt, S., Coulston, R., & Lunsford, R. (2004). When do we interact multimodally? Cognitive load and multimodal communication partners. In *Proceedings of the ICMI'04* (pp. 129–136). New York: ACM Press.

Oviatt, S., DeAngeli, A., & Kuhn, K. (1997). Integration and synchronization of input modes during multimodal human-computer interaction. In *CHI'97 Proceedings* (pp. 415–422). New York: ACM Press.

Oviatt, S., & Kuhn, K. (1998). Referential features and linguistic indirection in multimodal language. In *Proceedings of the ICSLP'98* (pp. 2339–2342).

Oviatt, S., Lunsford, R., & Coulston, R. (2005). Individal differences in multimodal integration patterns: What are they and why do they exist? In *CHI'05 Proceedings* (pp. 241–249). New York: ACM Press.

Oviatt, S., & van Gent, R. (1996). Error resolution during multimodal human-computer interaction. In *Proceedings of the ICSLP'96* (pp. 204–207).

Palomares Perraut, R., & Gómez Camarero, C. (2014). An Overview of Information Literacy Competencies for Translators Using Web 2.0 Environments. *International Humanity Studies*, *1*(2), 40–54.

Parush, A. (2005). Speech-Based Interaction in Multitask Conditions: Impact of Prompt Modality. *Human Factors*, *47*(3), 591–597.

Parush, A. (2015). *Conceptual Design for Interactive Systems: Designing for Performance and User Experience*. Waltham: Morgan Kaufman.

Pausch, R., & Leatherby, J. H. (1991). An Empirical Study: Adding Voice Input to a Graphical Editor. *Journal of the American Voice Input/Output Society*, *9*(2), 55–66.

Pérez-Ortiz, J. A., Torregrosa, D., & Forcada, M. L. (2014). Black-box integration of heterogeneous bilingual resources into an interactive translation system. In *Workshop on Humans and Computer-assisted Translation* (pp. 57–65).

Pratt, S. (1990). L'importance de la traduction à vue pour l'enseignement de l'interprétation et de la traduction. In *La traduction, profession créative : Actes du XIIe congrès mondial de la FIT*. Belgrade.

Pym, A. (2005). Localization: On Its Nature, Virtues and Dangers. *Synaps*, *17*. Retrieved from http://usuaris.tinet.cat/apym/on-line/translation/Localization_bergen.doc

Pym, A. (2011). What technology does to translating. *Translation & Interpreting*, *3*(1), 1–9. Retrieved from http://www.trans-int.org/index.php/transint/article/viewFile/121/81

Rabiner, L. (1997). Applications of speech recognition in the area of telecommunications. Retrieved from http://www.ece.ucsb.edu/Faculty/Rabiner/ece259/Reprints/341_telecom applications.pdf

Reddy, A., & Rose, R. C. (2010). Integration of Statistical Models for Dictation of Document Translations in a Machine Aided Human Translation Task. *IEEE Transactions on Audio, Speech and Language Processing*, *18*(8), 1–11. Retrieved from http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=05393062

Rodriguez, L., Reddy, A., & Rose, R. (2012). Efficient Integration of Translation and Speech Models in Dictation Based Machine Aided Human Translation. In *Proceedings of the IEEE 2012 International Conference on Acoustics, Speech, and Signal Processing* (Vol. 2, pp. 4949–4952).

Romero, V., Leiva, L., Toselli, A. H., & Vidal, E. (2009). Interactive multimodal transcription of text images using a web-based demo system. In *Proc. IUI* (pp. 477–478).

Romero-Fresco, P. (2011). *Subtitling Through Speech Recognition: Respeaking*. Manchester: St. Jerome.

Romero-Fresco, P. (2012). Respeaking in Translator Training Curricula. Present and Future Prospects. *ITT*, *6*(1), 91–112.

Sáenz Zapata, L. M. (2011). Integration of Ergonomics in the Design Process: Conceptual, Methodological, and Practical Foundations. In W. Karwowski, M. Doares, & N. A. Stanton (Eds.), *Human Factors and Ergonomics in Consumer Product Design* (pp. 155–175). Boca Raton: CRC Press.

Sales, D. (2008). Towards a student-centred approach to information literacy learning: A focus group study on the information behaviour of translation and interpreting students. *Journal of Information Literacy*, *2*(1), 41–60.

Sales, D., & Pinto, M. (2011). The professional translator and information literacy: Perceptions and needs. *Journal of Librarianship and Information Science*, *42*, 246–260.

Schalkwyk, J., Beeferman, D., Byrne, B., Chelba, C., Cohen, M., Garret, M., & Strope, B. (2010). Google Search by Voice: A Case Study. In *Advances in Speech Recognition* (pp. 61–90). New York: Springer. Retrieved from http://static.googleusercontent.com/external_content/untrusted_dlcp/research.google.com/en/us/pubs/archive/36340.pdf

Seaman, L. (2002). Voice Recognition for Translators, Or Why I Started Talking To My Computer. *JLD Times. Newsletter of the Japanese Language Division of the American Translators Association*, 6–8.

Shilman, M., Tan, D. S., & Simard, P. (2006). CueTIP: a mixed-initiative interface for correcting handwriting errors. In *Proc. UIST* (pp. 323–332).

Shneiderman, B. (1998). *Designing the User Interface: Strategies for Effective Human-Computer Interaction* (3rd ed.). Reading, Massachusetts: Addison-Wesley.

Shneiderman, B. (2008). Foreword. In A. Sears & J. A. Jacko (Eds.), *The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications* (2nd ed., pp. xix–xx). New York: Lawrence Erlbaum Associates.

Shneiderman, B. (2015). Foreword. In A. Parush (Ed.), *Conceptual Design for Interactive Systems: Designing for Performance and User Experience* (pp. ix–x). Waltham: Morgan Kaufman.

Shneiderman, B., & Plaisant, C. (2010). *Designing the User Interface: Strategies for Effective Human-Computer Interaction* (5th ed.). Reading, Massachusetts: Addison-Wesley.

Simard, T. (2013). *Analyse comparative de la couverture et de l'acceptabilité des solutions d'un dictionnaire bilingue spécialisé, d'une banque de données terminologiques et d'un concordancier en ligne : application au domaine de la traduction médicale*. Université d'Ottawa. Retrieved from http://www.ruor.uottawa.ca/fr/handle/10393/24929

Spilka, I. V. (1966). La traduction à vue instrument de formation. *Meta*, *11*(2), 43–45. Retrieved from http://www.erudit.org/revue/meta/1966/v11/n2/010462ar.pdf

Stone, R. J. (2001). Haptic feedback: A brief history from telepresence to virtual reality. *Haptic Human-Computer Interaction*, 1–16. doi:10.1007/3-540-44589-7_1

Stroman, J. (2002). Translation and Voice Recognition Software. *JLD Times. Newsletter of the Japanese Language Division of the American Translators Association*, 1–5.

Swallow, D., Blythe, M., & Wright, P. (2005). Grounding experience: relating theory and method to evaluate the user experience of smartphones. In *Proceedings of the Annual conference on European association of cognitive ergonomics* (pp. 91–98).

Taravella, A. (2013). L'homme et la machine, duo productif… et positif ? *L'Actualité Langagière*, *10*(2). Retrieved from http://www.bt-tb.tpsgc-pwgsc.gc.ca/btb.php?lang=fra&cont=3076#art7

Taravella, A., & Villeneuve, A. (2011). Aspects humains des technologies langagières dans l'organisation. In *Actes de la conférence de Tralogy, Paris, le 3 mars 2011*. Retrieved from http://lodel.irevues.inist.fr/tralogy/index.php?id=134

Taravella, A., & Villeneuve, A. (2013). Acknowledging the needs of computer-assisted translation tools users: the human perspective in human-machine translation. *The Journal of Specialised Translation*, *19*(January 2013), 62–74. Retrieved from http://www.jostrans.org/issue19/art_taravella.pdf

Teixeira, C. (2014). Data collection methods for researching the interaction between translators and translation tools: An ecological approach. In J. Schweiter & A. Ferreira (Eds.), *The development of translation competence: Theories and methodologies from psycholinguistics and cognitive science.* (pp. 269–286). Newcastle upon Tyne.

Touré, M. A., Mbangwana, M., & Sène, P. A. (2009). Que sont les TIC : Typologie des outils et des systèmes. In T. Karsenti (Ed.), *Intégration pedagogique des TIC en Afrique. Stratégies d'action et pistes de réflexion* (pp. 33–56). Ottawa: CRDI.

Uebel, L. F., & Woodland, P. C. (1999). An investigation into vocal tract length normalisation. In *Sixth European Conference on Speech Communication and Technology* (pp. 1–4). Retrieved from http://www.isca-speech.org/archive/eurospeech_1999/e99_2527.html

Uebler, U. (2011). Multilingual speech recognition in seven languages. *Speech Communication*, *35*(1-2), 53–69.

Vicente, K. (1997). *Cognitive work analysis: Towards safe, productive, and healthy computer-based work*. Mahwan, New Jersey: Lawrence Erlbaum Associates.

Vicente, K. (2004). *Le facteur humain : réinventer notre rapport à la technologie* *(trad. N. Guillet)*. Outremont: Logiques.

Vidal, E., Casacuberta, F., Rodríguez, L., Civera, J., & Martínez Hinarejos, C. D. (2006). Computer-assisted translation using speech recognition. *IEEE Transactions on Audio, Speech and Language Processing*, *14*(3). Retrieved from http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=01621206

Volanen, S. (2015). *Translating with the Web: Professional translators' information-seeking behaviour in translation with online resources*. MA Thesis. University of Turku, Finland.

Wahlster, W. (1991). User and discourse models for multimodal communication. In J. W. Sullivan & S. Tyler (Eds.), *Intelligent User Interfaces* (pp. 45–67). New: ACM Press.

Weizenbaum, J. (1966). A computer program for the study of natural language communication between man and machine. *Communications of the Association for Computing Machinery*, *9*(1). Retrieved from http://delivery.acm.org/10.1145/370000/365168/p36-weizenabaum.pdf?ip=192.75.139.249&acc=ACTIVE SERVICE&CFID=112302521&CFTOKEN=88721827&__acm__=1339788366_6a0abf705088a4f74ca702c72b26059e

Williamson, J. R., Crossan, A., & Brewster, S. (2011). Multimodal Mobile Interactions : Usability Studies in Real World Settings. In *Proceedings of the 13th International Conference on Multimodal Interaction* (pp. 361–368).

Woodland, P. C. (2001). Speaker adaptation for continuous density HMMs: A review. In *ISCA Tutorial and Research Workshop (ITRW) on Adaptation Methods for Speech Recognition.*

Yu, S., Hu, S., Zhang, S., & Xu, B. (2003). Chinese-English Bilingual Speech Recognition. In *NLPKE´03 Proceedings* (pp. 603–609).

Zapata Rojas, J. (2012). *Traduction dictée interactive : intégrer la reconnaissance vocale à l'enseignement et à la pratique de la traduction professionnelle*. Université d'Ottawa. Retrieved from http://www.ruor.uottawa.ca/en/bitstream/handle/10393/23227/Zapata Rojas_Julian_2012_these.pdf?sequence=1

Zapata, J. (2016). Investigating Translator-Information Interaction: A Case Study on the Use of the Prototype Biconcordancer Tool Integrated in CASMACAT. In M. Carl, S. Bangalore, & M. Shaeffer (Eds.), *New Directions in Empirical Translation*

Process Research (pp. 135–152). New York: Springer.

Zapata, J., & Kirkedal, A. S. (2015). Assessing the Performance of Automatic Speech Recognition Systems When Used by Native and Non-Native Speakers of Three Major Languages in Dictation Workflows. In *Proceedings of the 20th Nordic Conference of Computational Linguistics (NODALIDA 2015)* (pp. 201–210).

Zhang, Q., Pan, J., & Yan, Y. (2008). Mandarin-English Bilingual Speech Recognition for Real World Music Retrieval. In *Proceedings of the ICASSP´08* (pp. 4253–4256).

Žibert, J., Martincic-Ipsic, S., Ipsic, I., & Mihelic, F. (2003). Bilingual Speech Recognition of Slovenian and Croatian Weather Forecasts. In *Proceedings of the 4th EURASIP Conference* (pp. 637–642).

Zuckerberg, M. (2014). Connecting the World from the Sky. Retrieved from https://fbcdn-dragon-a.akamaihd.net/hphotos-ak-ash3/t39.2365-6/851574_611544752265540_1262758947_n.pdf

APPENDIX A — University of Ottawa Research Ethics Board Approval Notices for Main Experiment

File Number: 03-14-09 **Date (mm/dd/yyyy):** 04/16/2014

Université d'Ottawa
Bureau d'éthique et d'intégrité de la recherche

University of Ottawa
Office of Research Ethics and Integrity

Ethics Approval Notice

Social Science and Humanities REB

Principal Investigator / Supervisor / Co-investigator(s) / Student(s)

First Name	Last Name	Affiliation	Role
Jean	Quirion	Arts / Translation	Supervisor
Julian	Zapata Rojas	Arts / Translation	Student Researcher

File Number: 03-14-09

Type of Project: PhD Thesis

Title: Integrating the Human Factor in Translation Tool Design: A Comparative Analysis of Various Translators' User Experience with Multimodal Interfaces

Approval Date (mm/dd/yyyy)	Expiry Date (mm/dd/yyyy)	Approval Type
04/16/2014	04/15/2015	Ia

(Ia: Approval, Ib: Approval for initial stage only)

Special Conditions / Comments:
N/A

File Number: 03-14-09 **Date (mm/dd/yyyy):** 04/16/2014

Université d'Ottawa University of Ottawa

Bureau d'éthique et d'intégrité de la recherche Office of Research Ethics and Integrity

This is to confirm that the University of Ottawa Research Ethics Board identified above, which operates in accordance with the Tri-Council Policy Statement and other applicable laws and regulations in Ontario, has examined and approved the application for ethical approval for the above named research project as of the Ethics Approval Date indicated for the period above and subject to the conditions listed the section above entitled "Special Conditions / Comments".

During the course of the study the protocol may not be modified without prior written approval from the REB except when necessary to remove participants from immediate endangerment or when the modification(s) pertain to only administrative or logistical components of the study (e.g. change of telephone number). Investigators must also promptly alert the REB of any changes which increase the risk to participant(s), any changes which considerably affect the conduct of the project, all unanticipated and harmful events that occur, and new information that may negatively affect the conduct of the project and safety of the participant(s). Modifications to the project, information/consent documentation, and/or recruitment documentation, should be submitted to this office for approval using the "Modification to research project" form available at: http://www.research.uottawa.ca/ethics/forms.html.

Please submit an annual status report to the Protocol Officer four weeks before the above-referenced expiry date to either close the file or request a renewal of ethics approval. This document can be found at: http://www.research.uottawa.ca/ethics/forms.html.

If you have any questions, please do not hesitate to contact the Ethics Office at [redacted] or by e-mail at: [redacted].

Signature:

[redacted]

Protocol Officer for Ethics in Research
For [redacted] Chair of the Social Sciences and Humanities REB

[redacted]

APPENDIX B — Experimental Material

Text Used for Pre-control Task

Nanotubes de carbone et photosynthèse

Le concept de nanobionique appliquée au monde végétal a germé dans le laboratoire du Pr. Strano alors qu'il souhaitait développer des cellules solaires capables de s'auto-réparer à l'instar de la plupart des plantes. Dans le même temps, son équipe a travaillé sur les technologies à base de nanotubes de carbone susceptibles d'augmenter la photosynthèse des plantes. En effet, les chloroplastes ne sont sensibles qu'à la partie visible du spectre, laquelle ne représente qu'environ 50 % de l'énergie incidente. Le Pr. Strano a eu l'idée d'incorporer des nanotubes de carbone au sein des chloroplastes afin de bénéficier de la large capacité d'absorption du spectre lumineux de ces nanoparticules; celles-ci sont capables de capter les longueurs d'onde non absorbées par les plantes, de l'ultraviolet à l'infra-rouge proche.

Text Used for Experimental Task

Bionic plants could light the future

MIT researchers harness the energy and potential of plants

Instead of mobile phone towers and street lights, imagine a world where trees light our way at night and branches are antennae.

In experiments on chloroplasts extracted from baby spinach leaves and thale cress, researchers from the MIT demonstrate that they can make plants sensitive to certain substances, opening the path to their use as sensors; boost the plants' photosynthesis activity so they are at least three times more effective than control plants; and enhance plant repair, allowing for the possible harnessing of plant energy.

Interview Guide – Control Group

Données sociodémographiques

1. *À quelle période de votre vie (petite enfance, enfance, adolescence, adulte) avez-vous acquis le français et l'anglais ?*

2. *Travaillez-vous indifféremment vers ces deux langues ?*

Situation actuelle

3. *Décrivez les étapes que vous suivez normalement pour produire une traduction.*

4. *Quels outils utilisez-vous normalement pour préparer, produire et réviser une traduction ? Pourquoi préférez-vous ces ressources ?*

5. *De 1 à 10, quel est votre niveau de satisfaction avec les techniques et les outils que vous utilisez couramment ? Pourquoi ?*
   Pas satisfait du tout ***1 2 3 4 5 6 7 8 9 10*** *Très satisfait*

Traduction dictée interactive

6. *Avez-vous déjà utilisé la reconnaissance vocale au travail ou dans votre quotidien ?* ***Sensation de productivité/qualité améliorée ?***

7. *Possédez-vous un téléphone intelligent ou une tablette ? Si oui, utilisez-vous un stylet comme dispositif de saisie ? Pourquoi ?*

8. *Possédez-vous un ordinateur à écran tactile ? Si oui, utilisez-vous encore le clavier et la souris de façon importante ? Pourquoi ?*

9. *Selon vous, qu'est-ce qui empêche la traduction dictée et la reconnaissance vocale d'être adoptées par un plus grand nombre de langagiers à l'heure actuelle ?*

Interview Guide – Test Group

Données sociodémographiques

1. *À quelle période de votre vie (petite enfance, enfance, adolescence, adulte) avez-vous acquis le français et l'anglais ?*

2. *Travaillez-vous indifféremment vers ces deux langues ?*

Situation actuelle

3. *Décrivez les étapes que vous suivez normalement pour produire une traduction.*

4. *Quels outils utilisez-vous normalement pour préparer, produire et réviser une traduction ? Pourquoi préférez-vous ces ressources ?*

5. *De 1 à 10, quel est votre niveau de satisfaction avec les techniques et les outils que vous utilisez couramment ? Pourquoi ?*
Pas satisfait du tout ***1 2 3 4 5 6 7 8 9 10*** *Très satisfait*

Traduction dictée interactive

6. *Aviez-vous déjà utilisé la reconnaissance vocale au travail ou dans votre quotidien ?*

7. *Comment vous êtes-vous senti en traduisant un texte par la voix ?*
Sensation de productivité/qualité améliorée ?

8. *Possédez-vous un téléphone intelligent ou une tablette ? Si oui, utilisez-vous un stylet comme dispositif de saisie ? Pourquoi ?*

9. *Possédez-vous un ordinateur à écran tactile ? Si oui, utilisez-vous encore le clavier et la souris de façon importante ? Pourquoi ?*

10. *Quelles difficultés avez-vous rencontrées avec les appareils, les dispositifs et les logiciels que vous avez utilisés pendant cette expérience ?*

11. *Selon vous, peut-on utiliser cette technique et ces outils pour traduire tout genre de textes ? Pourquoi ?*

12. *Selon vous, qu'est-ce qui empêche la traduction dictée et la reconnaissance vocale d'être adoptées par un plus grand nombre de langagiers à l'heure actuelle ?*

APPENDIX C — A Pilot Experiment on Source-Text Preparation in Translation Dictation: Methodology and Results

Methodology

Considering the scope of a pilot experiment, I limited data collection methods to in-person active observation and semi-structured interviews.

A call for participants was sent out to graduate students in the Master's program in conference interpretation or other graduate students who had received interpretation training in the past. I limited the sample size to four participants. The general objective of the experiment was to learn about TD workflows, specifically in CTDD. Specifically, I was interested in learning how translators/student interpreters prepared their source text prior to dictating the target text into a digital dictaphone.

Translators with a background in interpreting have a developed capacity to orally translate a text, and a certain degree of familiarity with various dictation tools. For this experiment, the focus was put on the source-text preparation step, which implied the use of an array of tools, from (online) informational resources (e.g. dictionaries, terminology databases and corpora) to (physical) writing tools (e.g. pencils, pens and highlighters).

A CTDD workstation was simulated for the experiment, as shown in Figure 1 below. The experiment took place in a computer lab at the University of Ottawa and was carried out individually (i.e., one participant at a time). The experimental environment was composed of a desktop PC connected to an Internet network, a digital dictaphone, a paper copy of the source text, a pencil, and eraser, a pencil sharpener, highlighters of four different colors (fuchsia, blue, yellow and orange) and pens of five different colors (fuchsia, blue, purple, green and black):

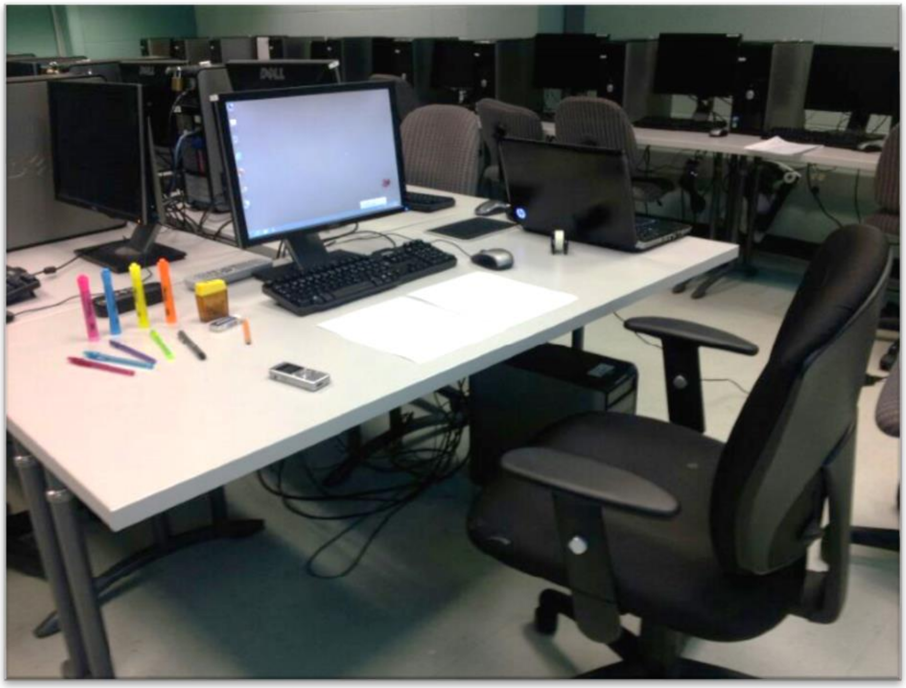


Figure 1. Experimental environment: A conventional translation dictation with dictaphone workstation

Participants were given a paper copy of a 300-word English text, printed double-spaced, and were instructed to translate it into French using the CTDD technique. The source text was taken from a technology news website. The text was intentionally chosen so that it would be fairly easy to sight translate, yet difficult enough for the translator to use the various (online and physical) tools to prepare for the dictation.

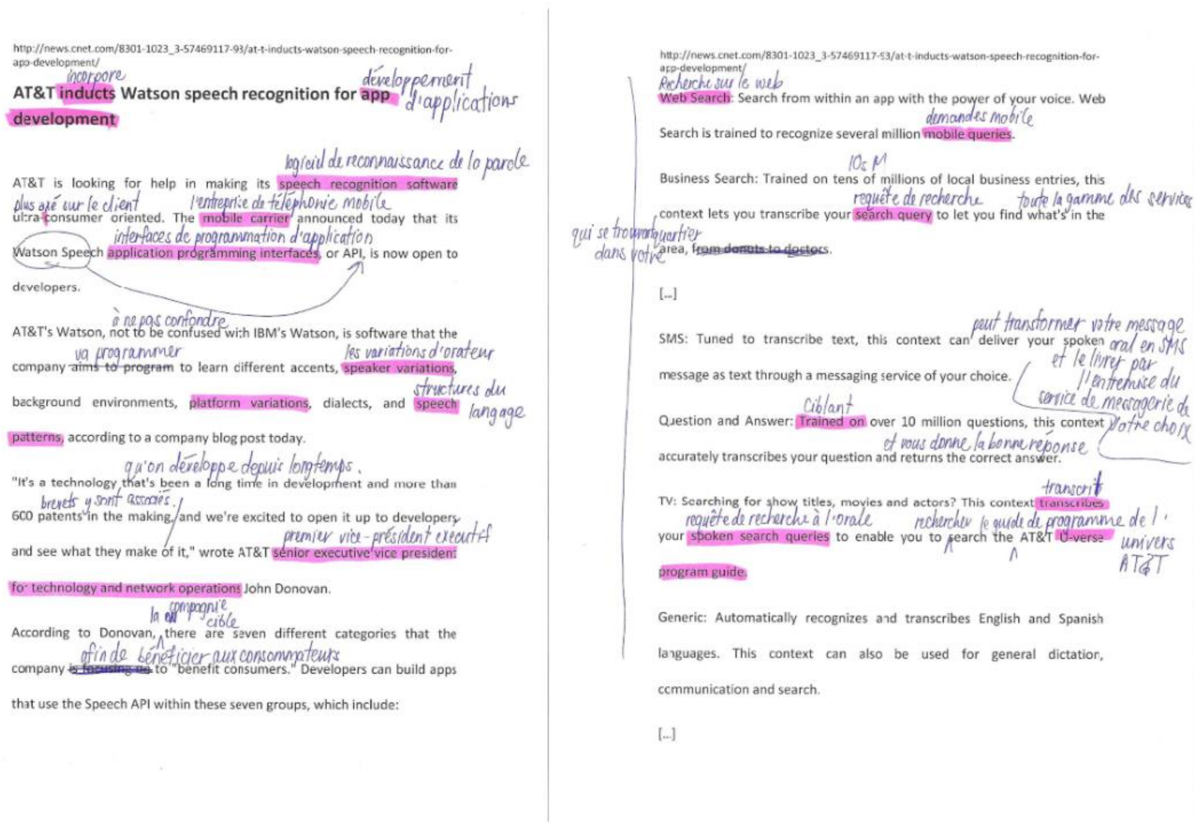
Each participant was allowed 15 to 30 minutes to prepare their translation using all the tools made available to them. Following the source text preparation step, participants recorded their translation using a digital dictaphone. By the end of the exercise, participants were instructed to hand in the source text copy containing their annotations. Lastly, during a short interview, participants responded to a few socio-demographic questions (e.g. about their age group and their linguistic and educational background) as well as certain questions about the process of preparing and producing a TD.

In order to maximize participants' anonymity (since it is a small sample and most of participants know each other) special care was taken for storing participants' data. For instance, the codes used here to present the data (P1, P2, P3 and P4) do not correspond to the order in which participants carried out the experiment. In addition—and for the same reasons—no link is made between the data and the participant's gender, age group or linguistic background. These criteria, however, may be essential when analyzing data from a larger-scale experiment and thus including a greater number of participants.

Experimental results

The sample included two male and two female graduate students between the ages of 24 and 39 years old. Only one person reported to have French (the target language in this experiment) as their mother tongue. The rest of the participants where translating either towards their second or their third language. None of the participants acquired the experiments' two languages (English and French) at early childhood. Two participants reported to be right-handed and two left-handed. Out of four participants, one affirms to have been trained in oral translation techniques many years ago (SiT, TD), while the remaining three received such training only recently. Lastly, one participant says to have worked professionally in TD, while the remaining three have only used it as a technique it in simultaneous interpretation training.

Source text preparation tasks have been under-investigated in TS, including how translators annotate source text copies prior to producing the target texts. There would be no "right" or "wrong" way to prepare a source text; it is rather an intuitive action, which depends on the translator's level of experience and/or level of comfort with the source and target languages, and on the availability of tools. Figure 2 shows an example of a source-text copy annotated by one of the four participants:


http://news.cnet.com/8301-1023_3-57469117-93/at-t-inducts-watson-speech-recognition-for-app-development/

incorpore — développement d'applications

AT&T inducts Watson speech recognition for app development

logiciel de reconnaissance de la parole

AT&T is looking for help in making its speech recognition software ultra-consumer oriented. The mobile carrier announced today that its Watson Speech application programming interfaces, or API, is now open to developers.

plus axé sur le client — l'entreprise de téléphonie mobile — interfaces de programmation d'application

AT&T's Watson, not to be confused with IBM's Watson, is software that the company ~~aims to program~~ to learn different accents, speaker variations, background environments, platform variations, dialects, and speech patterns, according to a company blog post today.

à ne pas confondre — va programmer — les variations d'orateur — structures du langage

"It's a technology that's been a long time in development and more than 600 patents in the making, and we're excited to open it up to developers and see what they make of it," wrote AT&T senior executive vice president for technology and network operations John Donovan.

qu'on développe depuis longtemps. — brevets y sont associés. — premier vice-président exécutif

According to Donovan, there are seven different categories that the company ~~is focusing on~~ to "benefit consumers." Developers can build apps that use the Speech API within these seven groups, which include:

la compagnie cible — afin de bénéficier aux consommateurs

http://news.cnet.com/8301-1023_3-57469117-93/at-t-inducts-watson-speech-recognition-for-app-development/

Web Search: Search from within an app with the power of your voice. Web Search is trained to recognize several million mobile queries.

Recherche sur le web — demandes mobile

Business Search: Trained on tens of millions of local business entries, this context lets you transcribe your search query to let you find what's in the area, ~~from donuts to doctors~~.

10s M — requête de recherche — toute la gamme des services qui se trouvent dans votre quartier

[...]

SMS: Tuned to transcribe text, this context can deliver your spoken message as text through a messaging service of your choice.

peut transformer votre message oral en SMS et le livrer par l'entremise du service de messagerie de votre choix

Question and Answer: Trained on over 10 million questions, this context accurately transcribes your question and returns the correct answer.

Ciblant — et vous donne la bonne réponse

TV: Searching for show titles, movies and actors? This context transcribes your spoken search queries to enable you to search the AT&T U-verse program guide.

transcrit — requête de recherche à l'orale — rechercher le guide de programme de l'univers AT&T

Generic: Automatically recognizes and transcribes English and Spanish languages. This context can also be used for general dictation, communication and search.

[...]

Figure 2. A source text copy annotated by one of the participants

During the experiment, we observed the participant's screen activity and other behaviors. The participant's computer screen was connected to a projector, which allowed me to observe screen activities as they were performed. An alternative to this would have been using a screen recording application, but considering the scope of the experiment, I refrained from installing such an application in a University lab computer.

Figure 3 shows the time spent by participants reading and doing manual preparation before launching any search on the web.

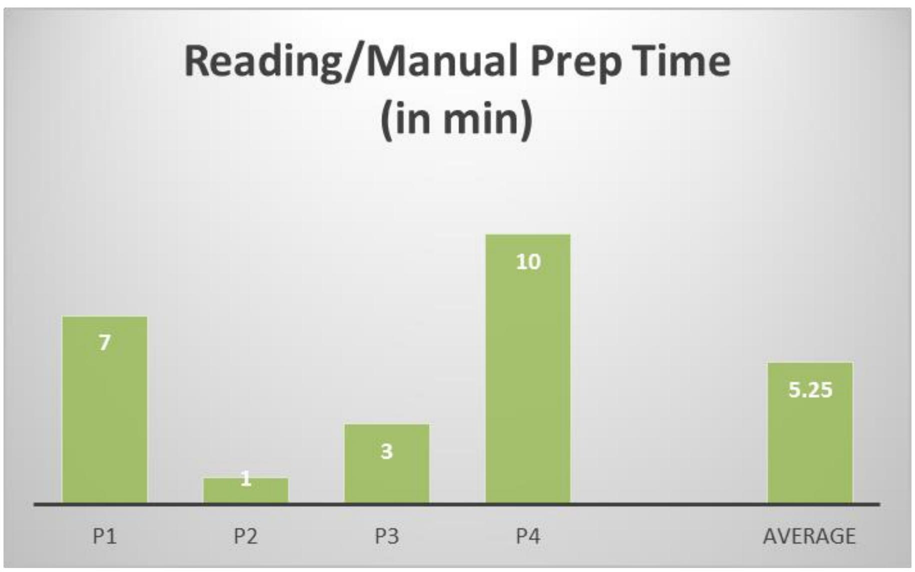


Figure 3. Reading and Manual Preparation time

Figure shows the time spent reading and doing manual preparation versus the total preparation time.

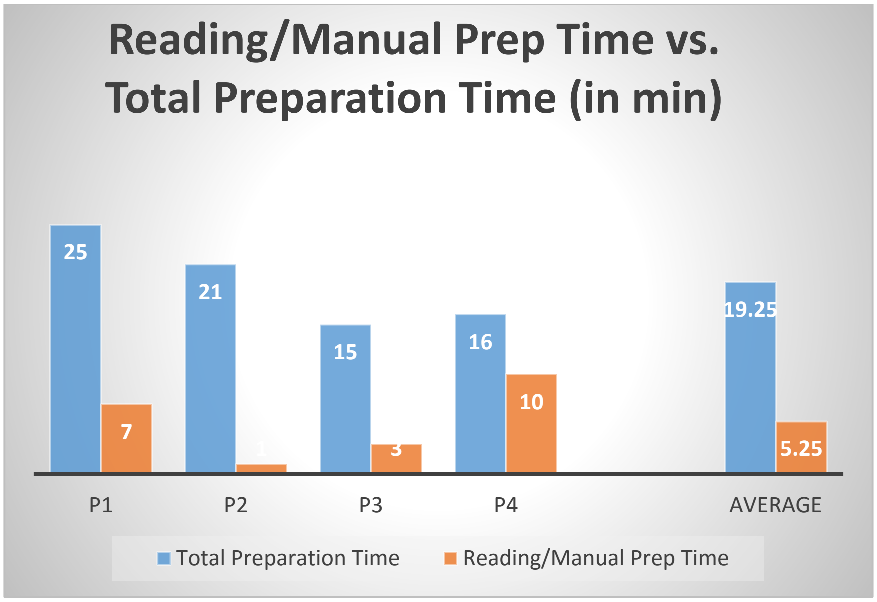


Figure 4. Reading and manual preparation time vs. total preparation time

Figure 5 shows the various online informational resources used and the percentage of participants who used each tool.

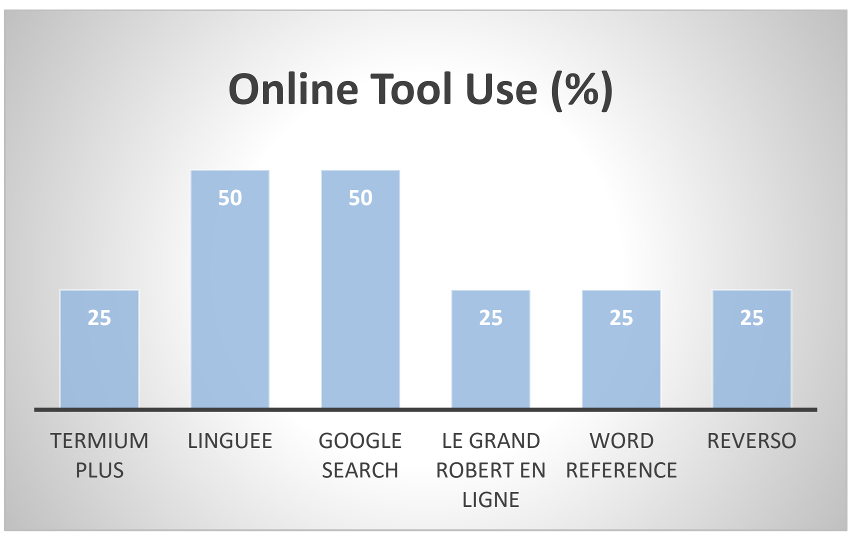


Figure 5. Online tool usage

Figure 6 shows the various writing tools used and the percentage of participants who used each tool.

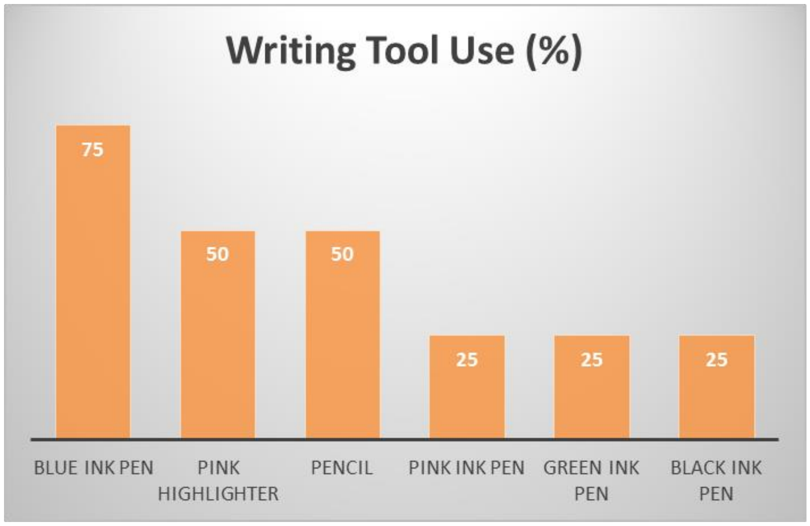


Figure 6. Writing tool usage

In addition, participants took part in an interview at the end of the exercise. The interviews lasted 35 minutes on average. Because it was a semi-structured interview, it was possible not only to obtain the answers to the pre-set questions but also to gain insight about some of the participant's specific behaviors and opinions. For instance, it was possible to ask participants why do they prefer a tool or another, why did they choose a certain color of pen or highlighter to prepare the source text, why would they use a pencil instead of a pen, and what their symbols and other annotations mean to them, and how does all this help them in order to dictate a translation. Some of their answers are compiled in the tables below.

Table 1 below compiles participants' verbalization of their source-text preparation behaviour:

"When preparing for a translation dictation, I…"
read the source text twice
highlight difficult terms
search online for terms in terminological databases
analyze the syntax and change the order of some sentences
highlight the length of paragraphs
highlight capital letters and punctuation marks
highlight acronyms
write symbols were possible (as in consecutive interpreting note-taking)
circle numbers and acronyms and names
circle terms I know already
sight-translate in my head

Table 1. Participants' verbalization of their source-text preparation behaviour

Table 2 compiles other source-text preparation and information behavior observed during the experiment.

A bilingual concordancer is used to search for words or phrases, as well as specialized terms
A search engine is used to search for definitions (e.g. by typing the word in English + "definition")
A search engine is used to search for equivalents in the target language (e.g. by typing the word in English + "*traduction*" in French)
A text processing application is used to write single words or sentences in both the source and target languages as reminders
During the dictation task, the recording is paused in order to search for a specific piece of information
All browser windows are closed after completing one single search, then open again
One does not necessarily read the entire source text prior to begin searching for information on the web

Table 2. Other source-text preparation behaviour noted during the experiment

Finally, table 3 compiles participant's verbalization of their use of the pen during the source-text preparation step as well as some examples taken from their writing samples:

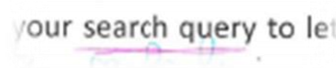
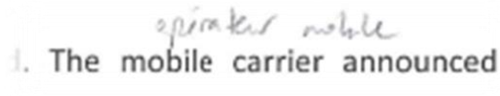
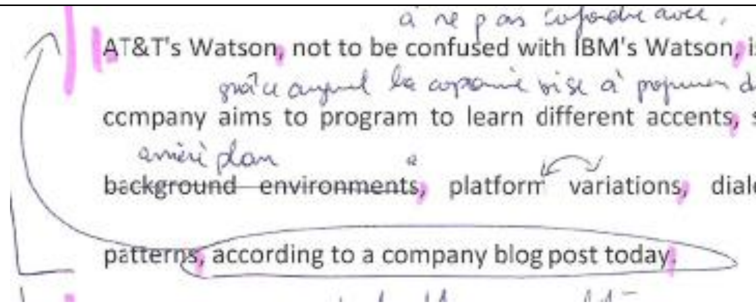
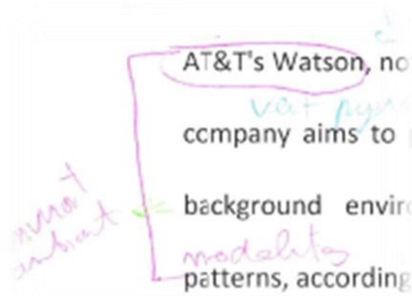
"I use the pen to…"	
underline difficult terms or passages	
write down equivalents	
make arrows for word reordering	
split sentences	
circle elements that would go in different places in the target text	
make annotations in margin	
scratch out non-essential stuff	

Table 3. Participants' verbalization of their use of the pen and examples from annotated source-text copies

APPENDIX D — University of Ottawa Research Ethics Board Approval Notices for Pilot Experiment on Source-Text Preparation for Translation Dictation

File Number: TRA6985 **Date (mm/dd/yyyy):** 12/21/2012

Université d'Ottawa
Bureau d'éthique et d'intégrité de la recherche

University of Ottawa
Office of Research Ethics and Integrity

Ethics Approval Notice

Social Science and Humanities REB

Principal Investigator / Supervisor / Co-investigator(s) / Student(s)

First Name	Last Name	Affiliation	Role
Elizabeth	Marshman	Arts / Others	Principal Investigator
Jean	Quirion	Arts / Translation	Co-Principal Investigator

File Number: TRA6985

Type of Project: Course Outline

Title: TRA6985: Current Developments in Translation Studies II

Approval Date (mm/dd/yyyy)	Expiry Date (mm/dd/yyyy)	Approval Type
01/01/2013	01/01/2014	Ia

(Ia: Approval, Ib: Approval for initial stage only)

Special Conditions / Comments:
N/A

Université d'Ottawa University of Ottawa

Bureau d'éthique et d'intégrité de la recherche Office of Research Ethics and Integrity

This is to confirm that the University of Ottawa Research Ethics Board identified above, which operates in accordance with the Tri-Council Policy Statement and other applicable laws and regulations in Ontario, has examined and approved the application for ethical approval for the above named research project as of the Ethics Approval Date indicated for the period above and subject to the conditions listed the section above entitled "Special Conditions / Comments".

During the course of the study the protocol may not be modified without prior written approval from the REB except when necessary to remove subjects from immediate endangerment or when the modification(s) pertain to only administrative or logistical components of the study (e.g. change of telephone number). Investigators must also promptly alert the REB of any changes which increase the risk to participant(s), any changes which considerably affect the conduct of the project, all unanticipated and harmful events that occur, and new information that may negatively affect the conduct of the project and safety of the participant(s). Modifications to the project, information/consent documentation, and/or recruitment documentation, should be submitted to this office for approval using the "Modification to research project" form available at: http://www.research.uottawa.ca/ethics/forms.html

Please submit an annual status report to the Protocol Officer four weeks before the above-referenced expiry date to either close the file or request a renewal of ethics approval. This document can be found at: http://www.research.uottawa.ca/ethics/forms.html

If you have any questions, please do not hesitate to contact the Ethics Office at [redacted] or by e-mail at: [redacted]

Signature:

[redacted]

Protocol Officer for Ethics in Research
Fo[redacted], Chair of the Social Sciences and Humanities REB

[redacted]