Tampere University

MARY NURMINEN

Investigating the Influence of Context in the Use and Reception of Raw Machine Translation

Tampere University Dissertations 515

Tampere University Dissertations 515

MARY NURMINEN

Investigating the Influence of Context in the Use and Reception of Raw Machine Translation

ACADEMIC DISSERTATION
To be presented, with the permission of
the Faculty of Information Technology and Communication Sciences
of Tampere University,
for public discussion in auditorium A1
of the university's main building, Kalevantie 4, Tampere,
on 3 December 2021, at 13 o'clock.

ACADEMIC DISSERTATION
Tampere University, Faculty of Information Technology and Communication Sciences
Finland

Responsible supervisor and Custos	Professor Kaisa Koskinen Tampere University Finland	
Supervisors	Professor (acting) Tiina Tuominen University of Turku Finland	Professor Kaisa Väänänen Tampere University Finland
Pre-examiners	Professor Sharon O'Brien Dublin City University Ireland	University Lecturer Leena Salmi University of Turku Finland
Opponent	Professor Lynne Bowker University of Ottawa Canada	

Cover design: Roihu Inc.

ISBN 978-952-03-2198-7 (print)
ISBN 978-952-03-2199-4 (pdf)
ISSN 2489-9860 (print)
ISSN 2490-0028 (pdf)
http://urn.fi/URN:ISBN:978-952-03-2199-4

PunaMusta Oy – Yliopistopaino
Joensuu 2021

To Mom and Mummi, two of the wisest people I know

Äidille ja Mummille, kahdelle viisaimmalle tuntemalleni ihmiselle

ACKNOWLEDGEMENTS

In 2006 I read an article describing how people rely on raw, unedited machine translation to access product technical support in their own language and became fascinated. 15 years later, that fascination has morphed into this dissertation. The five years of this project have been an exhilarating ride.

A triumvirate of advisors guided me on this journey, and I thank them for the thought and effort they've put into the project. Kaisa Koskinen (a.k.a. Kaisa K), I don't know how such a busy person manages to somehow clear her table and make time for meeting me whenever needed. That has been instrumental at so many points over the past five years. I thank you for the thought and ideas you've put into my topic and writing, and also for the time put into dealing with my occasional crazy ideas (best quote: "I don't know whether to cheer you on or to try to save you from yourself.").

Tiina Tuominen (a.k.a., well, Tiina), when I first approached you and explained how my topic of machine translation was closely tied to your research focus on subtitle reception, you quickly got over the pesky small detail of the topics being from two utterly different worlds and saw the connection. You then graciously agreed to act as an advisor. I thank you for your insightful questions and for the thoroughness with which you read and commented on the various parts of the dissertation, and for the advice you generously gave throughout.

The day I sent a mail to you, Kaisa Väänänen (a.k.a. Kaisa V), asking if you might be interested in joining the advising team, will always stand out in my mind. I was a bit afraid to approach such a busy person and ask about working on a project in another (at that time anyway) university. I was astounded to get an almost immediate reply that you were interested in the topic and would be glad to join. Although you had little time allocated to this PhD, you left a visible imprint on the final result. Your well-timed comments and questions helped me define frameworks for analysis and ask the right questions. Thank you for that.

My most heartfelt gratitude goes to my colleague and dear friend Sari Hokkanen, without whom this dissertation would have more than occasional spots of alarmingly poor quality. You read every piece as I produced it, edited it all thoroughly, and pushed me relentlessly toward improvement. When necessary, even in the face of a stressed-out writer at the end of her wits, you were even willing to use those frightful words that only the most fearless of editors dare to utter: "It's still not clear." I will never be able to thank you enough for that work and for the way you were willing to dive into the world of machine translation reception. Our intellectual conversations have been a delight. However, I also want to thank you for the non-intellectual things: the lunches, the conferences, the dinners, the discussions about food (anyone noticing a pattern here?).

Another dear friend I'd like to thank is Kate Moore. You diligently walked me through the fine points of academic writing, multiple times actually. Then edited and corrected my work until that one glorious day when you read a piece of the dissertation and delivered those precious words: "I didn't have much to comment on." You kept up a constant flow of encouragement and helped make sure I had what I needed to work, including a spot to write. Far more importantly, you called, joked, and got me laughing time and again. Sari and Kate, I thank you most of all for your friendship, which I cherish. And one last note – yes, I *did* start a sentence with a numeral up there in the first paragraph 😊

Three very important collaborators made significant contributions to this dissertation. Niko Papula of Multilizer Oy has been an enthusiastic collaborator who shares my passion to continuously learn more about how, when, where, and why people use raw MT. I hope our collaboration will continue to lead to new discoveries. Taina Rönkkö quite unexpectedly introduced me to a fascinating world in which machine translation is an everyday tool. She then played an instrumental role in helping me set up ways to study that world. Maarit Koponen has been the most excellent of co-authors. When we both suggested the same topic for a book chapter, it was you who jumped in and suggested we join forces and write together. That was three joint articles ago. The easy kind of cooperation we have had is a rare luxury in working life. I have enjoyed working with you, watching your mind at work, and learning from you. Last spring when crisis struck in the middle of our writing project, you graciously took over all of the work so that I could focus on my family's needs. Thank you for that too. Finally, I would like to thank the professionals who have given me their precious time to act as informants in my studies.

I would like to thank the pre-examiners of the dissertation, Professor Sharon O'Brien from Dublin City University and University Lecturer Leena Salmi from the University of Turku. Your constructive feedback and useful comments were clear and valuable. They made the finalization of the dissertation summary smooth.

A number of colleagues have contributed to this work directly or indirectly. Marja Kivilehto, your enthusiasm and support for my work has been one of the things that has kept me going. I also thank you for a well-timed question which you might not even remember since it was at the beginning of this project: "Are you recording those interviews you are conducting for the MT Stories website?" When I said no, you asked simply, why not? It's easy to do and can safeguard your data. And it has done just that. Sirkku Latomaa, it has been so delightful to build and teach the Machine Translation and Post-editing course with you. During this fall's course you have put up with my distractedness and not-always-functioning brain, gently reminding me that the dissertation project is important too and it's OK to take time for it. Anu Viljanmaa and Arja Nurmi, you have given me insights into research and asked good questions. Tytti Suojanen, I will always be sad that you didn't get to see this day and celebrate with us. I miss you so much ♥

I am grateful to my fellow researchers, starting with Félix do Carmo, with whom I've had the best debates about the definition and nature of translation, and whether machine translation is translation at all (shameless ad: catch our panel at the EST2022 conference!). I am also grateful to my fellow students in the Doctoral Programme in Language Studies for the camaraderie and support. It's been lovely to share it with you! I thank Johnny Riquet for shaping the PhD seminar into a useful space for us to learn the ropes of research and explore our own topics. Finally, I thank Noora Jormakka for her curiosity and willingness to challenge ideas. It was a pleasure to work on the advising team for her Master's thesis.

I thank Tampere University for the courage they showed in hiring an outsider and for the generous support of that outsider's (ongoing) metamorphosis from businessperson to academic. I have been granted time and support for conferences, a summer school that played a key role in the early days of my PhD project, a half-year research leave, and a 2-year position as a full-time researcher. Without that support, this project would still be struggling along in its early stages.

I am also grateful for the encouragement I've gotten from a distance. My mom and sisters Linda and Bark have enthusiastically followed the project and kept me going. Liisa Penrose has listened to the highs and lows and persistently reassured and encouraged me. Mummi ja Pappa ovat tukeneet minua aina ja kaikessa. That has all meant so much to me.

Last but far from least, I want to thank my family. Emmi, you have been a role model in research and PhD work. Maggie, you have steadfastly encouraged me but also reminded me that work isn't everything. Jack, you have come up with technical solutions and patiently acted as my Excel teacher and support. Mikko, you saved me in a few different ways at the end of the project. Watching the growth of your families and welcoming new members has provided me with joy and a healthy grounding in real life during these years. Markku, more than anything else, it was your unwavering support that has made this project possible. You have been patient, understanding, and flexible. You have asked good questions – some about my research, others more practical ("Would it be time to quit working now?") And you never, ever had any doubts, even when I did. Thank you and I love you.

Nokia, 10 November 2021

Mary Nurminen

ABSTRACT

This dissertation is concerned with how people use raw, unedited machine translation (MT) for "gisting," or to access texts that are in languages they do not speak, with the aim of understanding as much of the text as is needed for a specific purpose. Technology improvements and the growth in popularity of free online MT tools over the past 15 years has led to an exponential expansion in the number of people using MT, arguably making it one of the most important recent innovations in human communication. However, despite the ubiquity and importance of the phenomenon, to date it has received little attention in research.

The aims of the dissertation were to analyze a limited number of contexts in which MT gisting takes place, to identify factors in those contexts that influence people's use and reception of raw MT, and finally, to examine theoretical frameworks that can help to conceptualize the phenomenon of MT gisting. The exploratory approach of the dissertation eventually led to four different contexts being analyzed, each of which was unique and required distinct methods of study. Most of the studies relied on qualitative methods; one study was quantitative. Results were published in five articles, and the dissertation comprises those articles plus this dissertation summary.

The four contexts studied were online MT, the use of raw MT in a professional ecosystem, MT-mediated interviewing for research, and the use of MT to increase accessibility to information. The study on the online MT context involved a survey of users of one online MT tool, and the results revealed a very diverse user group who relied on MT for a variety of purposes, the most prominent of which was gisting. Study was the most popular area of life respondents reported using MT in, followed by work and leisure. A somewhat surprising finding was that a large majority (83%) reported having some level of understanding of the language they were translating texts from, a possible indication that they use raw MT in a different way than they use human translations.

In the second context, patent professionals working in the intellectual property rights (IPR) field used raw MT on a regular basis to understand patent documents that are in languages they do not speak. MT gisting in this context was shaped by the fact that it occurred inside an ecosystem that accommodated it. The general riskiness

of IPR work provided affordances and a tolerance for the risk involved in using raw MT, and the practice of using raw MT was considered legitimate and supported. A second finding was that patent professionals employed a process of risk assessment and management in evaluating when and where to rely on raw MT. Finally, the use of raw MT in the context was analyzed through the concept of distributed cognition, and the concept was found to be an appropriate way to analyze and understand the practice.

The third context involved data-gathering interviews for research which were conducted via MT-mediated communication, meaning that the interviewer and participants communicated through a chat application with integrated MT. The interviewer and participants did not share a language; rather, each typed messages in their own language and those messages were machine-translated for the other participant. The study concluded that the method was promising and deserving of further study. It also uncovered seven considerations for using MT-mediated communications.

The final context concerned situations in which groups of people lack accessibility to the information they would need in order to participate fully in society, primarily due to a lack of competence in the languages in which information is published. The dissertation reviewed research and projects that aimed to increase accessibility to information and consequently, to improve access to civic life, health and safety information, and culture and media. It also explored promising new developments as well as challenges, including ethical ones, that are involved when using MT for increasing accessibility.

The second aim of the dissertation was to identify factors in the contexts studied that influenced people's use and reception of raw MT. Eleven contextual influences were identified and classified as relating to users, their tasks and goals, or the technical and organizational environment. User qualities that influence MT gisting included users' competences in the source or target languages, their familiarity with the textual context of the texts they were machine-translating, and their level of MT literacy. Factors that related to user tasks and goals included a tendency to access multimodal information in source texts to augment understanding of the raw MT, a tendency to verify raw MT output by using multiple MT tools, a tendency to adapt input to produce better MT output, and a tendency to negotiate the meaning of raw MT with others. Environmental aspects that affected MT gisting involved auxiliary technologies that augmented MT use, the status and legitimacy MT enjoys in the environment, and affordances provided by an environment that is accustomed to negotiating the meaning of texts and is tolerant of risk.

Finally, three different frameworks through which MT gisting can be conceptualized were proposed. The phenomenon can be viewed and analyzed through the framework of contextual influences, it can be conceptualized as an exercise in risk management, or it can be viewed through the framework of distributed cognition. It was proposed that the frameworks can be applied in both academic and industrial settings, for example, in future research, in evaluations of the suitability of raw MT for various use cases, in product or process development, or in the development of MT literacy programs.

The dissertation contributes new insights into some previously under-researched or unexplored contexts in which MT gisting is occurring. Its highlights and elevates the role of context in the use and reception of raw MT. Finally, it offers new alternatives for viewing and analyzing the phenomenon of MT gisting.

TIIVISTELMÄ

Tässä väitöskirjassa tarkastellaan editoimattomien konekäännösten käyttöä sellaisenaan tilanteissa, joissa vieraskielisestä tekstistä pyritään ymmärtämään mahdollisimman paljon tiettyä tarkoitusta varten. Englanniksi tällaista raakakonekäännösten käyttöä kutsutaan tutkimuksessa termillä ”gisting”. Teknologian kehittyminen ja internetissä vapaasti saatavilla olevien konekääntimien käytön yleistyminen viimeisten 15 vuoden aikana on johtanut konekääntämisen käyttäjämäärien räjähdysmäiseen kasvuun. Voidaankin sanoa, että konekääntäminen lienee yksi merkittävimmistä viimeaikaisista innovaatioista ihmisten välisessä viestinnässä. Ilmiön yleisyydestä ja tärkeydestä huolimatta sitä on kuitenkin tutkittu vasta vähän.

Väitöskirjan tavoitteena oli analysoida, miten raakakonekäännöksiä käytetään sellaisenaan erilaisissa konteksteissa, tunnistaa niitä tekijöitä, jotka näissä konteksteissa vaikuttavat raakakonekäännösten käyttöön ja vastaanottoon, ja lopuksi tarkastella teoreettisia viitekehyksiä, joiden avulla raakakonekäännösten käyttöä voidaan käsitteellistää ilmiönä. Väitöstutkimuksen kartoittava tutkimusote johti lopulta neljän kontekstin tarkempaan analyysiin. Nämä neljä kontekstia olivat keskenään erilaisia ja vaativat toisistaan poikkeavia tutkimusmenetelmiä. Useimmissa osatutkimuksissa käytettiin laadullisia menetelmiä; yhdessä osatutkimuksessa menetelmät olivat määrällisiä. Tutkimustulokset julkaistiin viidessä artikkelissa, jotka sisältyvät väitöskirjaan tämän yhteenvedon kanssa.

Tutkitut neljä kontekstia olivat verkkopohjainen konekääntäminen, raakakonekäännösten käyttö ammatillisessa ekosysteemissä, konekäännösvälitteinen tutkimushaastattelu ja konekääntämisen käyttö saavutettavuuden parantamiseksi. Verkkopohjaista konekääntämistä tarkastelevassa osatutkimuksessa laadittiin kysely, joka suunnattiin erään verkkopohjaisen konekääntimen käyttäjille. Tulokset osoittivat, että konekääntimen käyttäjäkunta oli erittäin monimuotoinen ja hyödynsi konekääntämistä monenlaisiin tarkoituksiin, joista yleisin oli väitöskirjan tarkastelema raakakonekäännösten käyttö sellaisenaan. Tyypillisin elämänalue, jolla vastaajat käyttivät konekääntämistä, oli opiskelu, toiseksi yleisin oli työ ja kolmanneksi yleisin vapaa-aika. Jokseenkin yllättäen osatutkimuksessa kävi ilmi, että suurin osa (83 %) vastaajista kertoi ymmärtävänsä konekäännetyn tekstin lähdekieltä edes jonkin verran, mikä saattaa kertoa siitä, että he käyttävät konekäännöksiä eri tavalla kuin perinteisiä ihmisten laatimia käännöksiä.

Toisessa tutkitussa kontekstissa tarkasteltiin immateriaalioikeuksien (IPR) parissa työskenteleviä patenttiasiantuntijoita, jotka käyttivät raakakonekäännöksiä

säännöllisesti ymmärtääkseen vierailla kielillä laadittuja patenttiasiakirjoja. Tässä kontekstissa ympäröivä ekosysteemi muovasi ja tuki raakakonekäännösten käyttöä. Koska IPR-työhön ylipäänsä kuuluu merkittäviä riskejä, ekosysteemi tarjosi toimijoille tilaa ottaa huomioon myös raakakonekäännösten käyttöön liittyviä riskejä ja antoi tukea niiden sietämiseen. Raakakonekäännösten käyttöä pidettiin hyväksyttävänä menettelytapana ja sitä myös tuettiin. Osatutkimuksen toinen keskeinen löydös oli se, että patenttiasiantuntijat hyödynsivät omaehtoisia riskienarviointi- ja riskienhallintaprosesseja, joiden avulla he arvioivat, milloin ja missä tilanteessa raakakonekäännökseen voidaan turvautua. Osatutkimuksessa analysoitiin myös raakakonekäännösten käyttöä hajautetun kognition käsitteen avulla, mikä osoittautui hyödylliseksi lähestymistavaksi toiminnan ymmärtämiseen.

Kolmas konteksti käsitti aineiston keräämiseen tähtääviä tutkimushaastatteluja, jotka käytiin konekäännösvälitteisesti: haastattelija ja osallistujat siis kommunikoivat chat-sovelluksen avulla, johon oli integroitu konekäännin. Haastattelija ja osallistujat eivät puhuneet samaa kieltä, vaan he kirjoittivat viestejä omalla kielellään, jotka sovelluksen sisäinen konekäännin käänsi vastaanottajan kielelle. Menetelmä osoittautui lupaavaksi ja jatkotutkimuksen arvoiseksi. Lisäksi osatutkimuksessa tunnistettiin seitsemän seikkaa, jotka on syytä ottaa huomioon konekäännösvälitteisessä viestinnässä.

Viimeinen väitöstutkimuksessa tarkasteltu konteksti koski puutteita tiedon saavutettavuudessa, jonka vuoksi tietyt ihmisryhmät eivät voi osallistua yhteiskunnan toimintaan täysipainoisesti, erityisesti siitä syystä, että heillä ei ole riittävää osaamista tiedon julkaisukielessä. Tässä osatutkimuksessa tehtiin katsaus saavutettavuutta käsitteleviin tutkimuksiin ja hankkeisiin, jotka pyrkivät helpottamaan yhteiskunnan toimintaan osallistumista, terveyteen ja turvallisuuteen liittyvän tiedon saavutettavuutta sekä kulttuurin ja median saavutettavuutta. Tutkimuksessa kartoitettiin lupaavia uusia edistysaskeleita sekä eettisiä ja muita haasteita, joita liittyy konekääntämisen käyttöön saavutettavuuden parantamiseksi.

Väitöskirjan toinen tavoite oli tunnistaa niitä tekijöitä tutkituissa konteksteissa, jotka vaikuttivat raakakonekäännösten käyttöön ja vastaanottoon. Tutkimuksessa tunnistettiin yksitoista tällaista kontekstuaalista tekijää, ja ne jaoteltiin kolmeen ryhmään: käyttäjiin, käyttäjien tehtäviin ja tavoitteisiin sekä tekniseen ympäristöön ja organisaatioympäristöön liittyviin tekijöihin. Raakakonekäännösten käyttöön vaikuttaviin käyttäjien ominaisuuksiin lukeutui osaaminen lähde- ja kohdekielissä, perehtyneisyys konekäännettävän tekstin tekstikontekstiin sekä konekäännöslukutaidon taso. Käyttäjien tehtäviin ja tavoitteisiin liittyviin tekijöihin lukeutui taipumus täydentää raakakonekäännöksestä saatavaa tietoa lähdetekstissä

olevalla multimodaalisella tiedolla, taipumus vahvistaa raakakonekäännöksen tuotos muilla konekäännöstyökaluilla, taipumus muokata lähdetekstiä paremman konekäännöksen saamiseksi sekä taipumus neuvotella raakakonekäännöksen merkitysestä muiden kanssa. Ympäristöön liittyviin tekijöihin lukeutui konekäännösten käyttöä helpottavat aputyökalut, konekääntämisen nauttima status ja legitimiteetti kyseisessä ympäristössä sekä tekstien merkityksen neuvotteluun tottuneen ja riskejä sietävän ympäristön tarjoamat affordanssit.

Lopuksi väitöskirjassa esitetään kolme viitekehystä raakakonekäännösten käytön käsitteellistämiseksi. Ilmiötä voidaan tarkastella ja analysoida kontekstuaalisten vaikutteiden viitekehyksen avulla, se voidaan käsitteellistää riskienhallintatoiminnaksi tai se voidaan nähdä hajautetun kognition viitekehyksen kautta. Väitöskirjassa esitetään näiden viitekehysten soveltamista sekä tieteellisessä tutkimuksessa että yritystoiminnassa, esimerkiksi raakakonekääntämisen soveltuvuuden arvioinnissa eri tilanteisiin, tuote- tai prosessikehityksessä tai konekäännöslukutaiton kehittämiseen tähtäävien ohjelmien kehittämisessä.

Väitöskirja tuottaa uutta tietoa konteksteista, joissa raakakonekäännöksiä käytetään sellaisenaan ja joita on aiemmin tutkittu vain vähän, jos lainkaan. Tutkimus korostaa kontekstin tärkeyttä raakakonekäännösten käytössä ja vastaanottamisessa. Se tarjoaa myös vaihtoehtoisia tapoja tarkastella ja analysoida raakakonekäännösten käyttöä ilmiönä.

CONTENTS

ABBREVIATIONS

CL	Controlled Language
FPE	full post-editing
HT	human translation
IPR	intellectual property rights
LPE	light post-editing
MT	machine translation
PE	post-editing
ST	source text
TAUS	Translation Automation User Society

ORIGINAL ARTICLES

This thesis is comprised of this summary and the following original publications, reproduced here by permission. The original publications are referred to as Articles I through V in the summary.

Article I — Nurminen, Mary and Niko Papula. 2018. "Gist MT Users: A Snapshot of the Use and Users of One Online MT Tool." *Proceedings of the 21st Annual Conference of the European Association for Machine Translation:* 199–208.

Article II — Nurminen, Mary. 2019. "Decision-making, Risk, and Gist Machine Translation in the Work of Patent Professionals." *Proceedings of the 8th Workshop on Patent and Scientific Literature Translation*: 32–42.

Article III — Nurminen, Mary. 2020. "Raw machine translation use by patent professionals: A case of distributed cognition." *Translation, Cognition and Behavior* 3 (1): 100–121.

Article IV — Nurminen, Mary. 2016. "Machine Translation-Mediated Interviewing as a Method for Gathering Data in Qualitative Research: A Pilot Project." *New Horizons in Translation Research and Education* 4 (16): 66–84. Joensuu, Finland: University of Eastern Finland.

Article V — Nurminen, Mary and Maarit Koponen. 2020. "Machine translation and fair access to information." *Translation Spaces* 9 (1): 150–169.

1 INTRODUCTION

One billion users.

This is the best estimate we have of the number of people who use raw machine translation (MT). It is a rough estimate based on a few statistics given by some of the largest MT providers. In April 2021, Google announced that one billion downloads of the Android version of the Google Translate app had been downloaded (Pitman 2021), double the 500 million users they reported in 2016 (Turovsky 2016). And in 2016 Facebook claimed that 800 million people see their machine translations each month (Way 2018, 162). If anything, the estimate should be considered consertative, since Google's report covers only people who use Google Translate through the Android version of the application, not all users, and since Google and Facebook are only two of the many suppliers of MT. On the other hand, a claim that people *see* translations does not equate with them being users, so the number of Facebook users might be smaller. However, despite the roughness of the definition of *user* and these additional assumptions, one billion is perhaps the best estimate we can make of the magnitude of MT usage today.

Users of MT can be divided into two main groups. The first group consists of professional translators who take MT output and edit it until it reaches an acceptable level of quality. This type of editing is referred to as *post-editing,* and the translators who perform it are commonly referred to as *post-editors*. This group can be considered one of the most important groups of MT users, as they engage with raw MT on a regular basis and become familiar with the technology, the quality levels it is capable of, and the mistakes it most commonly makes. Professional post-editors are also important contributors to the development of MT. For example, they often act as evaluators of linguistic quality in MT development (Castilho et al. 2018, 23). However, despite their importance, post-editors comprise only a very small percentage of all users of MT. Pym (2012, 137–140) estimated that there were 333,000 professional translators globally. If we assume that all of these translators also do post-editing (which they currently do not) and compare that to the estimate

of 1 billion users altogether, then professional post-editors would comprise a mere .03% of MT users.

This dissertation is concerned with the second group of MT users, which is comprised of people who consume raw, unedited MT (hereafter referred to as *raw MT*) in order to understand a text that is in another language. This estimated 99.97% of users have a myriad of purposes for using raw MT and a large range of contexts in which they use it. It includes everything from people who type individual words into Google Translate for fun, to people who use raw MT in high-risk decision-making in professional contexts, to people who conduct conversations via MT with people with whom they do not share a language. This type of raw MT use is termed *MT gisting* in this dissertation (see Section 2.1 for a more thorough discussion).

Despite the significant number of people who use raw MT, to date, only a small body of research has been devoted to studying these users. A positive development over the past 15 years or so has been a steadily growing research interest in professional post-editors' use of MT (see Koponen et al. 2021, 3). Unfortunately, a similar growth in interest on the other users of MT has not occurred. Since the first study on MT gisting by Henisz-Dostert (1979), new studies have trickled rather than surged into the body of research on the topic. This lack of focus on the group has been commented on by various researchers over the years. For example, in his 2007 dissertation on free online MT, Gaspari (2007, 19) stated that, "in spite of [MT services in the internet] being popular resources that help to cater for the needs for translation and multilingual communication of a diverse and large population of users, to date very little research has been carried out in this area." Nine years later, Castilho claimed that, "little is known about how *end users* engage with raw machine translated text or post-edited text, or how usable this text is" (Castilho 2016, 2). Now in 2021, although the use of MT for gisting has grown exponentially, we continue to lack knowledge on this ubiquitous phenomenon.

1.1 The dissertation and motivations

This dissertation investigates the contexts in which people use raw MT to understand texts in other languages or to communicate with people with whom they do not share a language. The research explores MT gisting in four different contexts, including in online MT, in a business ecosystem, MT-mediated interviewing as a data-gathering technique in research, and in situations in which readers suffering from a lack of accessibility to information can gain such access through MT. The

dissertation identifies and describes 11 contextual factors that were observed or reported to influence people's use and reception of raw MT. Finally, the research proposes three ways of conceptualizing the phenomenon of MT gisting.

The dissertation is comprised of five academic articles, published between 2016 and 2020 and presented in the list of original publications at the beginning of this document, and this dissertation summary. The summary has two purposes. First, it reviews the results of the five articles. Second, it synthesizes the results and uses them to propose new definitions, hypotheses, and conceptualizations for the phenomenon of MT gisting. Results are intended to both contribute to the academic study of MT gisting and offer useful insights to organizations that already use, or are considering beginning to use, raw MT.

A number of motivations catalyzed the research for the dissertation. The first one was triggered ten years before the work began, while I was employed as a documentation and localization manager at a large multinational corporation. I read an article by Jaap van der Meer (2006) that introduced the idea of using FAUT, or fully automatic useful translation (ibid., 89) to enable the distribution of published information to greater numbers of people. I realized for the first time that MT could be used for more than post-editing and that it was a promising method for increasing information flow to new groups of people. My fascination with that idea lingered on, and when I joined academia several years later, I declared MT gisting as my main research interest.

A second motivation for the dissertation arose soon after that declaration, when I discovered how little research existed on MT gisting. In MT research at the time, the majority of focus was given to MT technologies and "approaches and issues directly related in different ways to the linguistic performance of MT systems, with an emphasis on the quality of the output in the target language produced by such systems" (Gaspari 2007, 20). A small but growing body of research existed on how translators used MT in post-editing, but non-translator users of raw MT were largely ignored. With the promising use cases introduced by van der Meer and a few others, and a quickly growing use of online MT, the need for new research was clear.

A final motivation was the conviction that, to better understand MT gisting and eventually provide better solutions for users, we need to study the contexts in which MT gisting occurs. Existing research had approached MT gisting by surveying actual users (see Section 3.2.1); by testing the reception of specific texts translated by humans, MT or MT with post-editing (see Section 3.2.2), with participants who were not actual but potential users of MT; or by testing the effects of individual factors on raw MT reception, mostly in MT-mediated communication (see Section 3.1).

There were few deeper qualitative studies, so most of the knowledge on context had been gained through surveys. My own conviction on the importance of studying context was catalyzed by Gaspari's (2007) Ph.D. thesis and recipient evaluation studies carried out by Bowker (2009; 2007), and is summed up by Gaspari (2007, 21) as follows: "the linguistic quality of the output provided by online MT services represents one of the factors that can be taken into account […] [but] other non-linguistic aspects are also deemed to be of great importance in serving the needs of Internet users and meeting their expectations."

1.2 Objectives and research questions

I had three goals for this dissertation. First, I wanted to know more about the contexts in which people are using raw MT. For example, where are they using it? For what purposes? What types of processes are they using MT gisting in? Second, I was interested in how the contexts in which people use raw MT influence their use and their reception of it. Third, I was interested in exploring possible ways to conceptualize or understand the phenomenon of MT gisting. The questions that guided my research are listed below and then discussed.

RQ1	For each of the contexts investigated in the dissertation: What is the nature of MT gisting in this context?
RQ2	What contextual factors influence the use and reception of raw MT?
RQ3	What theoretical frameworks might help us to conceptualize the phenomenon of MT gisting?

Research question 1 was framed broadly to accommodate my goal of exploring the contexts in which MT gisting takes place. At the beginning of the work, little knowledge existed of specific contexts and real-life applications of MT gisting and we needed information on all aspects of MT use. A broadly framed question would help accomplish that. Also, the contexts studied in the dissertation would vary greatly. For these reasons, I decided to use a fairly open research question that would accommodate a variety of use cases.

Research question 2 At the beginning of the work that led to this dissertation, I had a preliminary idea that specific factors in the contexts in which people use raw MT influence how they use and understand it. The influence of context on all types

of communication had been examined in previous research and was a generally accepted fact. However, context's influence on MT gisting had yet to be explored, and I wanted to investigate the question holistically but also identify specific factors that influenced use and reception. This preliminary idea of influencing factors helped me to define the topic and scope of this dissertation and was chronologically the first research question I had. However, to answer that question, I would first need to explore the nature of MT gisting in various contexts, and I therefore placed the question regarding influencing factors as research question two.

Research question 3 concerned the exploration of theoretical frameworks through which the phenomenon of MT gisting can be analyzed and understood. Besides the descriptive information that would result from questions 1 and 2, I wanted to consider possible conceptualizations for the phenomenon of MT gisting. This was important for not only the dissertation, but as a contribution to the literature on MT gisting, which suffered from a scarcity of theoretical development.

Table 1. presents the research questions that were addressed in each of the dissertation articles.

Table 1. Research questions addressed in each article

Article	Title	RQ1	RQ2	RQ3
I	Gist MT Users: A Snapshot of the Use and Users of One Online MT Tool	X	X	
II	Decision-making, Risk, and Machine Translation in the Work of Patent Professionals	X	X	X
III	Raw Machine Translation Use by Patent Professionals: A Case of Distributed Cognition	X	X	X
IV	Machine Translation-Mediated Interviewing as a Method for Gathering Data in Qualitative Research: A Pilot Project	X	X	
V	Machine Translation and Fair Access to Information	X	X	

1.3 Contributions

The results of this dissertation contribute to the growing body of research on the use and users of raw MT. It offers insights into some contexts for MT gisting, including online MT use, the use of raw MT in a business ecosystem, the potential use of MT gisting for interviewing for research, and in public service contexts in which MT could help groups of people achieve better access to information they

need. Some of these contexts have been under-researched in the past while others have gone unrecognized previous to the articles published as part of this dissertation.

This research provides evidence that specific factors in the context of use of MT gisting influence how people use and receive raw MT. Through the examination of factors uncovered in the dissertation and a compiled view to their influence on MT gisting, the dissertation highlights the role of context in the phenomenon. An exploration of this role has been missing from research on the use of raw MT.

In addition, the dissertation offers theoretical contributions on how to conceptualize the phenomenon of MT gisting, as well as development of the terminology used by the field. Finally, this work suggests the methodology of MT-mediated interviewing for data gathering, which would allow researchers to include more diverse groups of informants into their projects without addition costs.

The dissertation also makes contributions to the discipline of Translation Studies. It offers a new viewpoint to studies on the reception of machine-translated texts, which might bring insights into the reception of texts translated through other methods. It also contributes to the areas of workplace studies and studies on cognition and translation. Both of these areas have traditionally focused on the translation process, and this dissertation's focus on reception offers new understanding and insight.

1.4 Contents and structure of the dissertation summary

The first half of this summary, Chapters 1–4, describe the past research and concepts the dissertation drew from and how the research was conducted. Chapter 2 presents the key concepts the dissertation relied on. Chapter 3 contains a review of relevant literature. In Chapter 4 I describe the five articles that make up the dissertation as well as the data and methods that were used in the research. The second half, comprising Chapters 5–8, examines the results and conclusions of the dissertation. Chapters 5–7 present results related to each of the three research questions and Chapter 8 concentrates on conclusions and future research.

2 KEY CONCEPTS

This dissertation relied on four independent key concepts which other scholars have defined in their previous work: MT gisting, context of use, categories of MT use, and MT literacy. The first concept, MT gisting, has been used in the literature for a short period of time and its definition is still being developed. In fact, my own use of the term developed over the course of the dissertation work and in this chapter, I propose a new definition. The second concept, context of use, was defined in ISO standard 9241-210:2019 (ISO 2019) and was directly applicable in my dissertation. I did not develop it further although I did use it in various stages of the dissertation. The categories of MT use proposed by Hutchins (2010) were applicable with some alterations that made the model more intuitive while also drawing a clearer distinction between using MT to understand texts in foreign languages (or gisting) and using MT as a starting point for post-editing. The fourth concept is MT literacy, which was initially defined for the context of academic communications by Bowker and Buitrago Ciro (2019). With their work as a starting point, I propose a broader definition that is not context-specific but applies to all MT use.

2.1 MT gisting

One part of this dissertation was a process of development for the term *MT gisting*. I began with the term and definitions offered in past research (see Forcada 2010, 217; Doherty and O'Brien 2014, 40; and Bowker and Buitrago Ciro 2019, 23), used various forms of the term in different articles, and finally arrived at a new definition while writing this dissertation summary. Considering that the study of MT gisting is an emerging area of research and terminology tends to be in a state of fluctuation during such periods, this process was to be expected and is deserving of further discussion.

I offer my definition of MT gisting below, followed by a discussion of the development behind its key components, according to the order they appear in the definition.

> MT gisting is the process of knowingly consuming raw machine translation with the aim of understanding as much of its meaning as needed for a specific purpose.

Knowingly consumes. It is my position that a defining characteristic of MT gisting is that people who are engaged in it are aware that the information they are reading was translated by a machine and has not been edited by a human. They therefore approach the raw MT in a different way than they would approach texts that were fully produced by humans. A person might read raw MT without knowing that it is raw MT, but in that case, they are not engaged in MT gisting.

Raw machine translation. This refers to MT output that has not been edited by a post-editor or any other human. It has been discussed under several different names in past work, the most common of which have been *raw MT* and *unedited MT*. Another moniker, *gist MT,* was my own choice in Articles II and III (Nurminen and Papula 2018; Nurminen 2019) and has also been used by Martindale (2020). In some commercial settings, raw MT is referred to as *pure MT* and that name has been included in the names of services[1] and products[2]. Another term is FAUT or Fully Automated Useful Translation. The term was coined by van der Meer (2006) as a reaction to the goal of early efforts in MT research, namely, fully automatic high-quality translation (FAHQT). Bar-Hillel (1964) declared FAHQT a "nonfeasible" goal and van der Meer (2006) claimed that FAUT better reflected how actual users of the technology view it. The term was also used in Nurminen (2018b). In the final two articles of the dissertation as well as in this summary, I decided to use *raw MT* because it is simple and unambiguous.

Aim of understanding. An issue I struggled with throughout my research was the level of understanding implied by the word *gist*. This was concretized by a question from a fellow researcher posed in the early stages of my work: How much understanding does MT gisting entail? Does it mean that a person knows roughly what a text is about, that they understand the basic points of the text, or that they understand most of a text even though they do not understand every word? I deliberated that question throughout my dissertation work, plagued by the difficulty

[1] AdHoc translations: https://www.adhoc-translations.com/technology/machine-translation/

[2] Systran: https://www.systransoft.com/neural-machine-translation-nmt

of pinpointing a definite level of implied understanding and by the vagueness of the definitions offered by other researchers and even those I myself proposed.

Some indication of a level of understanding is included in several definitions of gisting and MT gisting, beginning with dictionary defintions of *gist* as pertaining to "the main point or part" of something[3] or "the most important pieces of information about something, or general information without details."[4] In definitions of *gist* and *gisting* in MT research, Forcada (2010, 217) referred to "an approximate idea of the content of a text," Doherty and O'Brien (2014, 2) to a "rough" understanding of a text, Bowker and Buitrago Ciro (2019, 23) to "the general idea of the meaning of a text." I myself defined the level of understanding involved in gisting as "basic" in Articles I and II (Nurminen and Papula 2018, 199; Nurminen 2019, 32) and "at least a basic level of understanding" in Article III (Nurminen 2020, 101).

After much consideration of the question and observations during my research, I decided that none of the definitions sufficiently described MT gisting. This is evidenced in Doherty and O'Brien (2014), which, after defining gisting as a "rough understanding," (ibid., 2) later noted that the raw MT output used in their study gave participants "more than just a *gist* of the meaning" (ibid., 49). In my research I observed that MT gisting covered a broad spectrum of understanding, from simply comprehending the topic to grasping almost all aspects of a text. I concluded that the level of understanding achieved through gisting fluctuates. Not only does the MT output quality fluctuate, contextual factors such as user qualities and environmental supports affecting how well a text is understood vary. Forcada (2010, 217) pointed to this fluctuation, arguing that the level of understanding that might be achieved depends not only on the technology and languages involved, but also on the user. I also concluded that fluctuation in understanding can occur within one text, with certain passages of a text being better understood than other passages. This all means that it would be almost impossible to know what level of understanding is achieved by specific users in specific contexts and for specific texts and subtexts. I therefore determined that a useful definition of MT gisting would be centered on the *aim* of understanding a machine-translated text and not the level of understanding that is actually achieved.

As much of its meaning as they need for a specific purpose. Another characteristic of MT gisting is that people are using it to find information they need for a specific purpose, and they concentrate on achieving the level of understanding needed for that purpose and not more.

---

[3] Merriam-Webster definition of *gist*: https://www.merriam-webster.com/dictionary/gist

[4] Cambridge Dictionary definition of *gist*: https://dictionary.cambridge.org/dictionary/english/gist

A final note in conclusion. It might be justifiable to differentiate between the term *MT gisting*, which focuses on the consumption of raw MT, and discussing *the use of raw MT*, which can include the consumption of raw MT as well as the use of the resulting information for some purpose. However, for the sake of simplicity, in this dissertation the two phrases are used interchangeably.

2.2 Context of use

When investigating a broad question such as how people use a technology or how they experience the reception of a translated text, one problem is simply knowing which questions to ask them. As noted by Tuominen (2012, 14), the goal and result of some reception research might be simply to "provide material for further, refined questions." In an under-researched area, past research can offer little help. Recognizing this problem, I decided to begin with a model that could provide a structure through which I could conceptualize the areas to be studied and to develop specific questions for surveys and interviews. That model was provided in ISO standard 9241-210:2019[5]: Human-centred design for interactive systems (ISO 2019).

The purpose of ISO 9241-210:2019 is to provide a framework through which people involved in designing interactive systems can analyze future scenarios in which the interactive systems will be used by humans. Such an analysis is an important element of a human-centred design process. My own goals were somewhat different than those of the target audience of ISO 9241-210. Whereas product designers analyze context to provide immediate input design processes for future products, I wanted to analyze context to generate new knowledge on how current systems are used. Despite this difference in goals, the framework provided in the standard was also appropriate for analyzing MT gisting contexts. I especially appreciated the stated goal of revealing "needs, problems and constraints that might otherwise be overlooked" (ISO 2019, 12), as it reflected my own goals in studying MT gisting contexts.

A key concept in the standard is the concept of *context of use*, or "[t]he characteristics of the users, tasks, and organizational, technical and physical environment" (ISO 2019, 12). The first important part of the definition of this

[5] As I conducted my dissertation work from 2016 to 2021, the version of ISO standard 9241-210 I relied on for most of the work was from 2010, not 2019. However, the 2019 version became available later in the project and the parts I used were nearly identical in the two versions. To maintain consistent referencing throughout this dissertation summary, I updated everything to reflect the 2019 version of the standard.

concept concerns the latter part of the term: context ***of use***. The study of machine translation and other technologies involves several contexts. The technology is created in a context of design or creation. People and companies evaluate the possible implementation of MT technology in a business or economic context. The content to be translated by an MT system is developed in a content creation context. The *context of use* usually occurs after these other contexts, when a person generates or takes the output of an MT system and does something with it. This might be a post-editor who edits the output for distribution, a person who uses the output to understand the original text, or two people who use the outputs of a system to converse with each other. *Context of use* should therefore be understood to specifically apply to situations in which a person is consuming MT output and using it for a specific purpose.

A detailed description of the three categories of context of use – users, goals and tasks, and environment – are summarized in Table 2. below.

Table 2. Summary of ISO 9241-210's categories of *context of use* (ISO 2019, 13)

Category	Summary of definition
Users	Various types of users and stakeholders. Relevant characteristics of users, including knowledge, skill, experience, education, training, physical attributes, habits, preferences, and capabilities
Goals & tasks	The ways in which users typically carry out tasks, the frequency and duration of performance, interdependencies and activities carried out in parallel with tasks. Also, risks associated with tasks
Environment	Technical environment: software, hardware, and materials. Relevant characteristics of the physical environment: thermal conditions, lighting, layout, and furniture. Relevant characteristics of the social and cultural environment: work practices, organizational structure, attitudes.

One aspect of the ISO standard's model deserves special attention because it might differ from views that Translation Studies scholars hold on the role of end users. In the ISO standard, users have multiple roles. First, as the humans that interact with products, they are assumed to be part of the event that the context of use surrounds. Second, as shown in the definitions given in Table 2. , the users and their characteristics are considered a part of the context of use. Third, the ISO standard emphasizes that users should be involved in design and development efforts since their involvement "provides a valuable source of knowledge about the context of use, the tasks, and how users are likely to work with the future product, system or

service" (ISO 2019, 6). They should therefore also act as informants that help describe and interpret the context of use to product designers.

In my own work, users are also part of the event that the context of use surrounds. That event, MT gisting, involves a human user reading raw MT output. The context might be understood to consist of elements that surround this event but not the elements of the event itself, meaning the user and the raw MT. Accordingly, the user would enter the context and be surrounded by it during the event. However, I decided to follow the ISO standard's paradigm in which the user is part of the event of MT gisting, but the characteristics that the user brings into the event – their background knowledge and experience – is also part of the context of use. Those characteristics affect the event similarly to the goals and tasks and environment, and therefore can justifiably be considered part of the context of use.

In the same vein, one could argue that the characteristics of the raw MT output should be considered part of the context. For example, the linguistic quality of the output is surely a factor that affects the event of MT gisting. Similarly, the ISO model would include characteristics of the future product to be part of the context of use. However, the ISO model does not include the product itself, perhaps because it does not yet exist or perhaps because it will be the recipient of the information gained from the analysis of the context of use. And in this issue, I again followed the ISO standard and did not include the raw MT output as an element of the context of use, but for different reasons. First, it would make the scope of the dissertation too broad. Second, and more importantly, there is already a great deal of research on MT linguistic quality, while the users of MT and the contexts in which they use it have received little attention. I therefore wanted to focus solely on context in this dissertation.

Besides using the concept of *context of use* at the beginning of the dissertation to help develop questions for surveys and interviews, I returned to it later. Throughout the research, I had collected information pertaining to RQ2: What factors in the contexts of people's use of raw MT influence their use and reception of it? I identified factors individually and began to systematically document them early in the process, but I did not follow a typology or model to search for, identify, or document them. It was a simple list that grew organically. Eventually the list was long enough to require a structured means for discussing and presenting it. I again found that the ISO categories of user characteristics, goals and tasks, and environment were appropriate for structuring and discussing the factors that were uncovered in this dissertation.

2.3 Categories of MT use

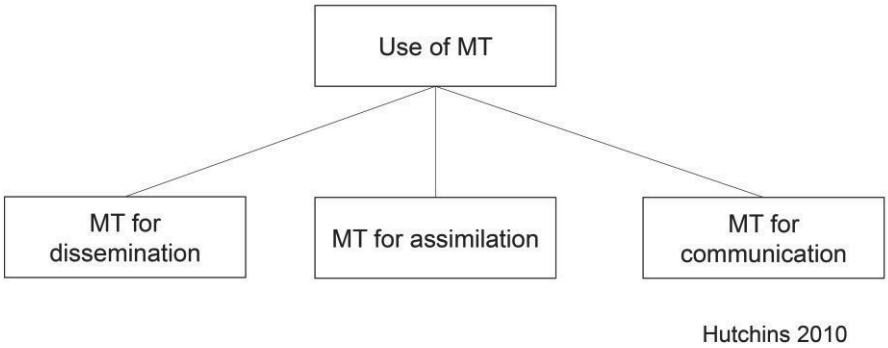
With the high number of people estimated to be engaged in MT gisting, it is clear that it is used in a wide variety of ways by a wide variety of different users. Categorizations by type of use or user can make it easier to discuss users and analyze their diverse needs. Despite this, surprisingly few proposals have been made on categories that would be useful for research and business. A few categorization models for MT use were proposed in the 1990s and 2000s, and researchers tend to continue to rely on them. In this section, I review the leading previous models for categorizing MT use and then propose a new one.

2.3.1 Previous categorizations of MT use

In their 1998 analysis of user feedback on the system named AltaVista Translation with Systran (commonly known as *Babelfish*), Yang and Lange (1998) identified five ways people used MT: as an assimilation tool, to comprehend information themselves for their own purposes; as a dissemination tool, for example, when people want to distribute the information on their own web pages to speakers of other languages by offering them MT; as a communication tool, so that they can communicate with people with whom they do not share a language; as an entertainment tool, such as when engaging in games with back-and-forth translations; and as a learning tool.

Hutchins (2010) listed three categories of MT use, and this three-category model has been cited widely in MT research. Hutchins used MT for dissemination to refer to a need for "translations of 'publishable' quality" (ibid., 29). This implied that a translator or post-editor would edit the MT output before publication. The next category, MT for assimilation, described the situation when a quick translation was needed which conveyed "the essence of the original text, however grammatically imperfect, lexically awkward and stylistically crude" (ibid.). MT for communication referred to cases in which MT was used for social interchanges such as e-mails or chatrooms. The categorization is illustrated in Figure 1. below.

Figure 1. Categories of MT use by Hutchins (2010, 29), which correspond to the three areas of MT user research that are currently most prominent



Koehn (2020, 19–28) included these three categories in addition to two further ones. He renamed MT for assimilation as information access and MT for dissemination as aiding human translators. Two new uses for MT were also included. First, MT can be used in natural language processing pipelines such as in search, and second, it can be part of a multimodal solution that includes both text translations and clues from visual elements. The final of these is an emerging area within MT.

In the early stages of my research, I primarily endeavored to work within the framework defined by Hutchins. However, through several realizations that came during the PhD project, I concluded that there were issues with the model that made it incompatible with the results of my own research. First, I found that MT for assimilation and communication shared important characteristics, indicating that they should be in the same category. Second, the differences between assimilation and communication on the one hand, and dissemination on the other, were not highlighted well enough. Finally, MT for dissemination was problematic in name and in how its assumed basis differs from that of the other categories. I discuss each of these realizations below.

When I began my dissertation work, it was clear to me that I wanted to analyze non-translators' use of MT. Relying on Hutchins's model, I assumed that this meant I needed to decide between focusing on MT for assimilation or on MT for communication. I had already published one article on MT for communication

(Article IV, Nurminen 2016), but was presented with an opportunity to study users of MT for assimilation and decided to pursue that choice and exclude the earlier article from the dissertation.

However, when collecting, analyzing, and reflecting on evidence for RQ2, concerning the contextual factors that influence the reception of raw MT, I came to the realization that the factors I had identified seemed to apply to MT for communication as well as MT for assimilation. For example, I had identified competence in the source language as a contextual factor that affected how people used and understood raw MT in assimilation scenarios. However, that same factor had also played a role in Article IV and appeared to have a similar influence. Eventually I found that 5 of the 11 factors identified applied to both assimilation and communication scenarios. Whereas certain contextual factors might apply more to one type than the other, all factors fit appropriately into one model, and the resulting model was more robust because both were included. It would seem that the two could justifiably belong to the same category. This realization led me to re-evaluate both the assumption that I needed to limit the scope to only one of the categories in Hutchins's model and my decision to not include Article IV in the dissertation.

The second issue that arose with Hutchins's model was that it did not highlight the differences between using MT for gisting (assimilation and communication) and using it for dissemination. This has led to a lack of focus on these differences, has potentially contributed to a siloed research field, and can make communication about research in the field difficult.

As described previously, my research pointed to similarities between MT for assimilation and communication. I also observed that as a group, these are inherently different than the category of MT for dissemination. This difference is predominately evident in the different responsibilities that users have in assimilation and communication scenarios as compared to dissemination scenarios. In assimilation and communication, users are primarily responsible for themselves and their own understanding of the received machine-translated text. Relying on the output as well as whatever contextual helps they have available (discussed in depth in Chapter 6), their main goal is to attain a sufficient understanding of the text for their own use. And as the best judges of what 'sufficient' understanding is, they only need to invest enough effort to achieve that and no more. In MT for communication, MT users might also assume the responsibility of producing texts that can easily be machine-translated for their conversation partners, but that is a separate process that occurs after their initial act of understanding their partners' most recent input.

The responsibilities are far greater for people editing raw MT in dissemination scenarios. Especially in the case of professional post-editing, post-editors are responsible to the original producers of the content to effectively reflect the meaning of that content. They are responsible to the commissioners of the post-editing work to fulfill the quality requirements that were agreed upon. Perhaps most importantly, they are also responsible to the readers of the post-edited content. Professional post-editors are not always aware of all of the contextual helps available to the eventual readers of the content. For example, they cannot assume all readers will be familiar with a specific genre or the culture that a text was created in. For this reason, professional post-editors need to produce a text that is understandable even when a minimum of contextual help is available. Besides translating meaning, they are responsible for knowing both source and target cultures and acting as cultural mediators. This responsibility can lead them to approach MT differently and focus on different issues than people using MT for gisting.

Hutchins's division into three types of use may have also contributed to a silo problem in the field of MT user research. The three categories reflect the three areas of MT user research that are currently most prominent: research on post-editing (MT for dissemination), MT gisting (MT for assimilation) and MT for communication, which have grown somewhat separately from each other. This is evidenced by a lack of citations between the areas and can, of course, be considered natural due to the need in research to focus on highly specific phenomena in well-delineated areas. However, the separation can also lead to a diminished sense of the phenomena as belonging to the same area of human communication and, more importantly, it can result in a reduction in the sharing of ideas and cooperation that might otherwise be possible.

Finally, in my own research, I noticed that the lack of a clear difference between the use of MT for post-editing and for other purposes could cause communication problems. When discussing my research with others, I noticed I often had to invest effort in explaining that the dissertation was *not* about translators but "the other users" of MT. This effort was needed whether I was explaining it to someone who knew nothing about translation or to someone who worked in Translation Studies. It was difficult to steer people away from the image of the professional translator as (the only) MT user, to an image of a different type of user. I found this curious, considering the large numbers of people using MT. If that many people were using raw MT, why was the image of the translator as the only user so engrained in people's minds? Even more curious was the fact that most of those people I was attempting to explain this to were also likely users of MT. There appeared to be a dissonance

between people's practice of using MT for various non-translator activities and their idea that this practice was worth discussing or studying. I often had the need to differentiate between two main categories of MT users, professional translators and non-translators, and Hutchins's model did not support such a differentiation.

The third issue with Hutchins's, and also Yang and Lange's, categorizations involved the category of MT for dissemination. Hutchins uses the term to describe situations in which "translations of 'publishable' quality" are required (Hutchins 2010, 29), a clear reference to post-editing. However, raw MT can also be disseminated; it can be passed from person to person or published. Actually, although they use the same term, dissemination, Yang and Lange (1998, 281) use it to specifically refer to the publishing of raw MT. This application of a term to refer to two different scenarios can cause confusion.

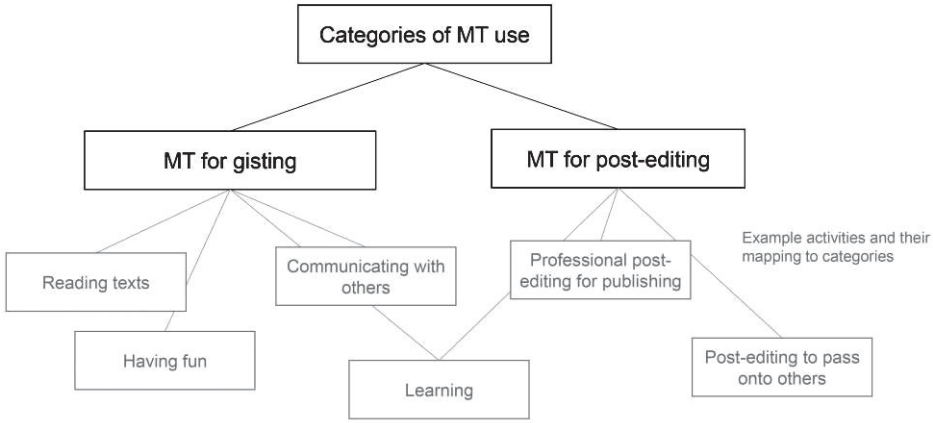
The category can also be problematic when considering what factors Hutchins's categorization is based on. If it is understood to be based on the actions taken on machine-translated information, then assimilation, communication and dissemination involve different levels of immediacy and different users. In MT for assimilation, machine-translated content is read by someone so that they can understand it; in MT for communication, it is read as input to an ongoing conversation; and in MT for dissemination, it is edited by a human so that it can be distributed further. The first two of these, assimilation and communication, reflect what happens *immediately* to the content by its initial user. The third, dissemination, describes an eventual activity, the distribution of the information, but this is neither immediate nor is it often undertaken by the initial user of the content, the post-editor. The content will often be further processed – for example, layout will be applied or perfected – and subsequently distributed by someone else through one or more channels. Even when the post-editor also acts as the distributor of the content, their first act would not be dissemination but post-editing.

After this process of realizing the issues with the model, I concluded that I needed to produce a model of categorizations for my own use. It is presented in the next section.

2.3.2 Categorization of MT use in this dissertation

Figure 2. presents the categorization model I developed and used in this dissertation. It is followed by a discussion outlining my reasons for the components of the model.

Figure 2. Categories of MT use adopted in this dissertation



The most important feature of this model is that it clearly delineates two main types of MT use, based on what users do with raw machine-translated content. In formulating a model with these two main categories of use, my first thought was to label them *professional* and *non-professional* use, but my research on patent professionals (Articles II and III) revealed that communities of users exist who are highly professional in their use of MT, even though they do not use it as a professional translator would. I therefore chose MT for gisting, which refers to the direct consumption of raw MT, and MT for post-editing, referring to situations in which raw MT is edited and corrected before being consumed by others. A further discussion of the wording of *MT for post-editing* follows later in this section.

The various ways people use MT can then be mapped to the categories. Examples of such activities are shown in gray in the model. These are meant to be just examples, not an exhaustive list of all activities in which people use MT. As research and our understanding of how MT is used increases, new examples will emerge.

As shown in Figure 2. some activities can be clearly mapped to one or the other of the categories. For example, people using MT to read a text in another language

or to communicate with someone else (Hutchins's MT for assimilation and MT for communication) are engaged in gisting. Using MT for having fun (Yang and Lange's MT use as an entertainment tool [1998, 201–202]), for example in Jimmy Fallon's Google Translate Songs feature[6], also involves the direct consumption of raw MT and would be classified as MT for gisting. On the other hand, both professional post-editing for publication (Hutchins's MT for dissemination) and the post-editing of a text to pass onto others in a more informal way, are mapped to the category of MT for post-editing.

It should also be noted that an activity might be mapped to both of the categories. The example in Figure 2. concerns the activity of *learning* (Yang and Lange's MT use as a learning tool [1998, 202–203]). It is clear that there are a variety of activities covered by learning. As pointed out by Jormakka (2021, 16), learning can encompass both post-editing exercises in language classes and using MT gisting for research. However, as research starts to unveil new knowledge about the myriad ways MT is and can be used in language and other learning, a better understanding of its place in categorizations of use may emerge.

Another feature of the model is the change from MT for dissemination to MT for post-editing. In addition to the issues with MT for dissemination that were discussed in Section 2.3.1, our knowledge of post-editing has increased and involved. Currently, post-editing is a commonly used term and the awareness of what it means and what it entails is greater than when the earlier models of MT use incorporated the concept of dissemination and therefore, the title MT for post-editing is more transparent.

I propose that this two-category solution resolves the issues I found with Hutchins's model, and that it can promote an understanding of the core differences in the different user groups' approaches to using MT. Through studies that explore each group separately as well as compare the differences between the two, we can gain an enhanced understanding of the needs and behavior of both groups.

[6] This feature involves machine-translating the lyrics of a popular tune back-and-forth between two distant languages, then singing the final result.
https://www.youtube.com/results?search_query=jimmy+fallon+google+translate+songs

2.4 MT literacy

A few of the early studies on MT use suggested that knowledge of MT and how it works might lead to an increased willingness to use MT and a better user experience with MT. In their 1993 article titled "Good Applications for Crummy Machine Translation", Church and Hovy (1993) advocated for finding suitable applications for MT, arguing that it would be better received if used in niche areas where it was determined to work well. They then defined six "desiderata" for a good niche application of MT. Several of these desiderata concerned technical or economic details, but two addressed the users' knowledge and expectations of the system: a good niche MT application "should set reasonable expectations" and "it should be clear to the users what the system can and cannot do" (Church & Hovy 1993, 256). In a similar vein, Miyazawa (2002, 9) encourage companies implementing MT to "[g]ive trainees a brief background of MT. Both strengths and weaknesses have to be communicated. This is a crucial part of the training. If this is not done, no one would be willing to use MT."

A promotion of understanding technology has also been suggested in relation to students' use of MT. For example, Williams (2006) introduced the need to include instruction on MT in language teaching and linked that instruction to electronic literacy. Bowker and Buitrago Ciro (2019) went further and coined the term *MT literacy*, then defined it for the specific context of scholarly communication as follows (Bowker and Buitrago Ciro 2019, 88):

> In the context of scholarly communication, machine translation literacy refers to a scholar's ability to:
>
> - comprehend the basics of how machine translation systems process texts;
> - understand how machine translation systems are or can be used (by oneself or by other scholars) to find, read, and/or produce scholarly publications;
> - appreciate the wider implications associated with the use of machine translation;
> - evaluate how (machine) translation-friendly a scholarly text is;
> - create or modify a scholarly text so that it could be translated more easily by a machine translation system; and
> - modify the output of a machine translation system to improve its accuracy and readability.

Adopting this definition as a base, Article II of this dissertation developed a second context-specific definition of MT literacy for the IPR context. As MT is used only for assimilation in that context, the definition (Nurminen 2019, 40) was shorter, referring to a patent professional's ability to:

> (1) comprehend the basics of how machine translation systems process texts, (2) understand how machine translation systems are or can be used (by oneself or others working with patents) to find and read patent documents within the context of IPR processes, and (3) appreciate the wider implications associated with the use of machine translation.

O'Brien and Ehrensberger-Dow (2020) acknowledge that Bowker and Buitrago Ciro's (2019) definition was intended for the specific context of academic research and publishing. However, they propose that it is applicable and useful in many other settings (O'Brien and Ehrensberger-Dow 2020, 146), including in translator training (ibid., 148), without changes.

2.4.1 Proposed general definition of MT literacy

Later in the dissertation work, I recognized the need for a definition of MT literacy that was not bound to specific contexts, so in this section I propose such a general definition. Bowker and Buitrago Ciro's definition offered a good basis for this because, although context-specific, it was broad enough to encompass the main aspects of MT use. In fact, their definition was broader than required here, as the scope of this dissertation is MT for gisting while their definition also included aspects of MT for post-editing. However, in the interest of working toward a generally applicable definition, the one proposed here includes those aspects as well.

MT literacy is defined by an MT user's ability to:

1. Comprehend the basics of how machine translation systems process texts
2. Understand machine translation systems' strengths and weaknesses
3. Understand how machine translation systems are or can be used for purposes that are important to the user
4. Appreciate the wider implications associated with the use of MT
5. Assimilate information from raw machine-translated texts
6. Evaluate how machine translation-friendly a text is
7. Create or modify a text so that it can be translated more easily by an MT system
8. Modify the output of an MT system to improve its accuracy and readability

Besides changing the wording to be more general, I also introduced two fundamental modifications to the definition proposed by Bowker and Buitrago Ciro, namely, the addition of points (2) and (5). A discussion of these changes follows.

Point 2, a user's ability to understand the strengths and weaknesses of machine translation systems, might be understood to be included in point 1 of Bowker and Buitrago Ciro's (2019) and Nurminen's (2019) definitions. I suggest the change because an understanding of strengths and weaknesses could be gained even without understanding how MT works, and there may be groups for whom that knowledge is enough. Also, evidence suggests that the understanding of strengths and weaknesses is sufficiently important that it warrants a separate point. In the patent professionals' context, it was important that they understood the capabilities of the systems they worked with. The study informants displayed an awareness of not only the capabilities but also the limitations of MT systems, and this influenced how they used MT and at which point they stopped using it and turned instead to human translation. As discussed in Articles III and IV, this is one of the most important decisions that patent professionals make in their work. The important competence of knowing what MT systems can and cannot do was also reflected in two of Church and Hovy's desiderata: a niche application "should set reasonable expectations" and "should be clear to the users what the system can and cannot do" (Church & Hovy 1993, 246). Viera, O'Hagan and O'Sullivan (2020, 1) also called for action on this point: "the review highlights a broad societal need for higher levels of awareness of the specific strengths and, crucially, of the limitations of MT."

The second addition I made was Point 5, the ability to assimilate information from raw machine-translated texts. This addition was based on discoveries I made

in this dissertation, many of which occurred during the study on patent professionals' MT gisting. In that study, informants described how newcomers were taught how to use MT tools but also how to read and interpret raw output. An excellent example of this came from informant PP4[7] who described something he commonly taught newcomers to his team on how to assimilate machine-translated information from patent documents that are relevant to the patent they are currently working on

> Typically, quite often the sequence of words appears to be more or less random because the Chinese grammar is not translated as well as the words themselves. So I advise people to be careful with words like *over* or *before* or *after* and look mainly at the nouns: which elements are present in the text. And then if...typically our interest is, are there any elements […] in the claim, in the patent text, which we do not have. Because if it is defined that there must be a seat, and there is no seat in our product, then we are free of that. So focusing on the nouns and if all the nouns match our product. Then try to decipher what is the relation between the nouns: what comes first and what comes second, etc…

Like the other competences of evaluating, creating, or modifying texts to be translated, and modifying the output of MT, the ability to assimilate information is not an absolute; a user does not either have the ability or not have it. Rather, users can demonstrate varying levels of competence in their understanding and assimilation of texts, and those abilities can develop through experience and education.

In conclusion, the use and further development of these concepts was necessary and instrumental during the dissertation process. However, it is hoped that it also contributes to ongoing and future research work in the area of MT gisting.

[7] PP4: Patent Professional 4, one of the informants of the study on patent professionals. In this and all subsequent quotes from informants, spoken data has been edited to conform to written language conventions.

3 PAST RESEARCH ON MT GISTING

This chapter provides a review of literature from various areas of MT research that discuss or analyze the use and users of MT, as well as research in related areas that provided methods, viewpoints or insights that were applied to the study of MT gisting in this dissertation. This review categorizes literature thematically, and within themes, chronologically from the oldest to the most recent. The first two sections cover research on two different types of MT gisting: MT for communication (Section 3.1) and MT for assimilation (Section 3.2). Although one proposal of this dissertation is that it would be beneficial to categorize these two as one group, MT for gisting (see Section 2.3.2), these areas have lived separate lives in the literature and therefore, this literature review discusses each separately. Section 3.3 presents a review of studies on MT use in various professional contexts. Section 3.4 examines research in which characteristics of appropriate use cases for raw MT have been proposed or analyzed. Sections 3.5–3.7 concentrate on research that provided a background for specific articles, with Section 3.5 offering a discussion of the literature on risk management in conjunction with MT gisting (Article II), Section 3.6 focusing on MT gisting and distributed cognition (Article III), and Section 3.7 discussing articles related to the use of MT to increase accessibility to information (Article V). Finally, literature in other related areas is introduced in Section 3.8.

3.1 MT for communication

MT for communication, also called MT-mediated communication, refers to situations in which raw MT is used to enable real-time or near real-time communication between people who do not share a language. Currently the most common implementation of MT-mediated communication involves an instant messaging or chat tool that is integrated with MT, so that each participant writes messages in their own language and those messages are translated for the other participant(s), who conversely reply with messages in their own languages which are then translated.

Research in MT for communication has focused predominantly on experimental studies. One interesting series of projects focused on intercultural collaboration, defined as "a goal-directed group activity" (Ishida 2016, 3–4). The first purpose of these projects was to develop and maintain an environment in which geographically and linguistically diverse groups could communicate and collaborate to achieve a goal, with machine translation as a key component in enabling that communication. A second purpose of the projects was to use this environment to observe participants' engagement with MT-mediated communication and to test the effects of various contextual factors on that communication. This focus on testing contextual factors was one reason that this group of studies was especially influential in this dissertation. After the initial project, the Intercultural Collaboration Experiment 2002 (ICE2002), ICE projects were conducted on an annual basis. Eventually the ICE work evolved into the Language Grid[8], a project that provides language resources which people can bundle to create their own environments in which they can employ intercultural collaboration to achieve a goal. From the beginning, the Language Grid has provided a platform for a number of intercultural projects, including YMC-Viet, which brought together agricultural experts from Japan and farmers in Vietnam to collaborate on achieving better agricultural results, and the Kyoto Intercultural Summer School for Youths (KISSY), an intercultural youth camp in which young people with different linguistic backgrounds collaborated through the support of MT[9]. As all communication in the projects occurred in MT-enabled chat systems, rich data in the form of transcripts was available for analysis immediately after the projects, or even during them. Numerous studies have been conducted since the beginning of the projects, including Nomura, Ishida and Yasouka (2003), Ogura et al. (2004), Yamashita and Ishida (2006a), Shigenobu (2007), Yamashita et al. (2009), Hautasaari (2010), Yasouka and Björn (2011), Suzuki and Hishiyama (2016), and Pituxcoosuvarn et al. (2018). A history of these projects can be found in Ishida (2016).

Many of the findings in experimental studies in the field of MT for communications have contributed to the current dissertation's exploration of contextual influences in MT gisting. They are discussed thoroughly in Chapter 6 and are therefore not described further in this section.

[8] Language Grid: https://langrid.org/en/index.html

[9] Language Grid projects: https://langrid.org/en/index.html#examples

3.2 MT for assimilation

An overview of research in the area of MT for assimilation reveals three types of studies and this section is structured accordingly. Studies based on surveys and experiments represent the majority of research in the area and are discussed in the first two sections. The final section is devoted to the far smaller category of qualitative research.

3.2.1 Survey studies

Since the beginning of MT, a number of surveys have been conducted on users and how they use MT. Some of these are from the early days of widely available, consumer-targeted MT and therefore take the form of market studies or reports on usage. Some studies focus on users of specific MT systems, while others study specific types of users regardless of the different MT systems involved. Recent user studies are components of projects developing MT solutions for specific purposes.

The International Association for Machine Translation (IAMT) sponsored research on users of various MT systems in 1993 and 1995, and these were reported by Lawson and Vasconcellos (1993), Vasconcellos (1994) and Brace, Vasconcellos and Miller (1995). The first of these had respondents from the Americas, Europe, and Japan, while the second focused on Europe and the Americas. Both studies concentrated heavily on MT use by professional translators, although "casual" users were included in minor ways, such as through "testimonials" (Lawson and Vasconcellos 1993, 127). These reports nevertheless also contained insights into the future importance of MT for gisting. For example, Brace, Vasconcellos and Miller (1995, 3) predicted that the potential impact of "on-line" MT on overall MT usage might be greater than its impact on MT as employed by professional translators.

Several surveys performed on MT users in Japan reflected a greater emphasis on the users of MT for assimilation as compared to professional translators in these early years of commercial MT. This was different than the situation in the U.S. and Europe. A 1995 survey studied the users of a Japanese commercial MT product, named Korya Eiwa, and reported that 92% of them were inexperienced with translation, implying that the great majority of the users of that product were engaged in MT for assimilation (Hoshino 1995, 2). The survey also asked respondents about their competence in English, their expectations for the system, their reasons for wanting to translate texts, and the types of texts they were translating with the

system. A few years later, the Asia-Pacific Association for Machine Translation (AAMT) conducted a series of surveys (2003, 2004 and 2005, reported on in Yamada et al. 2005) of internet MT users. This survey reported a high percentage of respondents who adopted MT for assimilation purposes and a low, but growing, adoption of MT for professional translation. The fact that the survey targeted internet MT users offers one explanation for this trend. Another explanation might be that, during that period, MT product manufacturers in Japan focused on assimilation users.

Three studies from the early- to mid-2000s reported on the implementation of MT for assimilation inside large international corporations. Morland (2002) described how NCR Corporation implemented MT and applied it to their company newsletter, then surveyed 280 of the people who had read the machine-translated versions of the newsletter. Morland's study found that respondents' perceptions of the usefulness of raw MT (or "pure" MT as Morland terms it) was higher than their rating of the translation quality, and that a high number of respondents (84%) reported it to be fairly useful or better (ibid., 7). In 2000, PricewaterhouseCoopers implemented an online MT function in their intranet that was available to all employees. In 2003, they conducted a survey of system users and uncovered an overall positive response to MT (Smith 2003). Nuutila (2005) reported on a survey of users of the Roughlate MT system implemented by Nokia in Finland in the mid-1990s. The purpose of the system was to allow employees to generate rough translations of texts in the form of raw MT. The user survey was conducted in 2002 and it found that both Finns and non-Finns were using the system, although it had been intended mainly for non-Finns. It also revealed that these two groups used the system differently, with Finns generating translations that they then edited and non-Finns generating translations for texts they needed to assimilate. Finally, the survey uncovered a positive perception on the usefulness of MT.

One of the most influential sources for this dissertation was the research conducted by Gaspari on users of online MT for assimilation. This included his PhD dissertation (Gaspari 2007) as well as other studies (Gaspari 2004; 2006; Gaspari and Hutchins 2007). One part of the dissertation was a survey which aimed to examine MT usage patterns among students at three UK universities. The survey first examined the types of people who were using MT, inquiring about demographics, languages translated, and informants' experience with computers and MT. It then focused on how the informants worked with MT, discovering that a large majority employed MT for assimilation or a combination of assimilation and dissemination, while few used it solely for dissemination. A large majority also reported translating

short texts, with a surprisingly high number translating only 1–2 words. This prompted Gaspari to include a section on the adoption of MT as an online dictionary, a type of use he referred to as not "entirely suitable" (Gaspari 2007, 106). An overview of research on MT for assimilation up to 2007 was produced by Gaspari and Hutchins (2007). They based the paper on a review of research as well as usage statistics they obtained from four online systems. They discovered that requests to translate texts, most often copy-pasted into MT tools, were more dominant than requests for translation of whole web pages. They also uncovered further evidence that people were translating very short texts – single words or short phrases – and proposed that this might be a sign that those people were working with languages they had some competence in, rather than languages they knew nothing of. The article also emphasized that, "The users of online MT are probably the largest group of MT users, and yet we know very little about them" (Gaspari and Hutchins 2007, 203).

Following an inactive period in research on users of MT for assimilation, a few more studies have emerged. Burgett (2015) reported on a survey conducted by Intel with visitors to their support site, many of whom accessed support documentation via raw MT. The goal of the paper was to examine these users' acceptance of receiving product information via raw MT. The survey questions concerned the tasks respondents performed with the help of support documentation, their success with completing those tasks, and also matters relating to language. A slight correlation was detected between the respondents' proficiency in English and a tendency for them to use both English and their own language while on the support site. A second user study was part of the Health in my Language (HimL[10]) project, which conducted two online studies of Polish and Romanian speakers in Scotland who were profiled as potential users of the eventual systems developed in the project. The two studies produced valuable information on the habits of people who access health information, specifically, that respondents reported encountering difficulties in accessing health information in their own languages, leading to them relying on Google Translate to translate information on health (Birch et al. 2017; 2018). Finally, in 2020, the European Commission conducted a survey of European small and medium businesses that represent the targeted users of their eTranslation tool (DG CNECT 2020). Of the 2,868 responses received, almost 40% reported that they did not currently use MT (ibid., 9). The top sectors represented by the responses were: manufacturing; professional, scientific and technical; other service activities; wholesale and retail trade; information and communication; and construction (ibid.,

[10] Health in my Language website: http://www.himl.eu/

5). A final finding involved respondents' views on suitable contexts and ways of using MT, including accessing it to understand websites or social media, gather information, look for business collaborators abroad, implement chatbots, or purchase products or services. Human translation was reported to be better in areas such as selling products or services, responding to bids, marketing, and legal negotiations and transactions (ibid., 8).

3.2.2 Experimental studies

A second type of study carried out on users of MT for assimilation were experimental studies that predominantly compared the reception of texts in two different ways. The first group made comparisons between different modes of translation – no translation (source texts or ST), human translation (HT), MT with no editing (raw MT), and MT with post-editing (full, FPE, or light, LPE). The second group made comparisons between the translations of texts that were authored through different methods, i.e. according to the rules of Controlled Language (CL) or not. A variety of testing methods and measures of reception were employed in these studies.

The largest group of studies used the testing of different types of reception factors to indicate which mode of translation led to better reception for end users. Each study compared 2–4 different translation modes, with the only mode that was included in every comparison being raw MT. The factors that were used as indicators of reception included comprehension, readability, confidence, quality, usability (comprised of goal completion, effectiveness, efficiency, and satisfaction), and acceptability (comprised of quality, usability, and satisfaction).

The first studies of this type concentrated on comparing source texts and raw MT. Fuji et al. (2001) tested the reception of ST, raw MT, and the combination of ST and raw MT on users with varying levels of proficiency in the source language. They found that showing end users both the ST and raw MT led to the best results for all types of users. Gaspari (2006) tested how well study participants could find basic information in web pages of a language they did not know, with one group working directly in the source language and the other using free online MT tools to access the text. The people using raw MT, unsurprisingly, scored higher on comprehension. However, they also showed higher confidence in the reliability of the web texts, which Gaspari interpreted as a positive perception of the usefulness of free online (raw) MT. Two studies by Doherty and O'Brien (2012; 2014) examined the usability of source texts (as consumed by users who knew the language)

compared to raw MT, as measured through eye tracking and questionnaires. In both studies, the ST performed better than raw MT. The 2012 article concluded by pondering whether two other modes, post-edited and human translation, should be included in such studies.

The idea of including human translation and post-editing to comparisons of STs and raw MT, which had been a part of two earlier studies (Bowker and Ehgoetz 2007; Bowker 2009), was soon taken up by others, and a number of subsequent studies tested the reception of texts produced through various translation modes. One study (Guerberof, O'Brien and Moorkens (2019) compared the reception of the ST, raw MT, and HT. The study found no significant differences in effectiveness and efficiency, but in satisfaction, HT scored significantly higher than raw MT. Another group of studies introduced PE to the mix, including Castilho et al. (2014), Bowker and Buitrago (2015), Castilho (2016), Castilho and O'Brien (2017), Hu, O'Brien, and Kenny (2019) and Girletti et al. (2019). The studies typically employed multiple methods of evaluation, including eye tracking, questionnaires, and preference ranking. In studies that only compared light or full post-editing and raw MT, post-editing led to a clear increase over raw MT in various usability measures. In studies that included raw MT, post-editing, and human translation, post-editing again performed positively.

An interesting point brought up in several articles (Doherty and O'Brien 2014; Castilho 2016; Castilho and O'Brien 2017; Hu, O'Brien and Kenny 2019) was that, even though tasks performed using raw MT might result in lower scores for the various factors tested, those tasks nevertheless were completed and raw MT was deemed usable. This was a good reminder that, although in comparison studies of this type, the use of raw MT might appear to be less efficient and effective than other modes of translation, that does not mean that it is not usable or useful. It also indicates that research focused on comparing raw MT to other modes of translation can only lead us to a certain amount of insight into the use of raw MT, and that different types of studies are also needed.

A second set of research made comparisons in the reception of texts that were authored without adherence to any sort of rules and texts that were authored according to various specific rule sets, e.g., CL rules. Both the source texts themselves and the texts resulting from the machine translation of the source texts were tested. Roturier (2006), O'Brien and Roturier (2007) and Doherty (2012) investigated the effect of specific rules of CL on the reception of raw MT alone, as tested through measures such as comprehensibility, usefulness, readability, and acceptability. O'Brien (2010) applied CL rules to a set of texts and examined recall,

readability, and acceptability of both the STs and the resulting raw MT versions of the texts. Collectively, these studies showed mixed results, with some CL rules leading to significant improvements and others resulting in smaller or no improvements. Bowker and Buitrago Ciro (2018) created versions of a ST through the application of two different rule sets: those derived from the rules of user experience and those that applied the rules of translatability. They had both versions machine-translated and fully post-edited, then evaluated: the ST by translators and all others by participants who represented assimilation users. The results revealed that the user experience rules, which led to a text which was more "engaging" and "catchy" (ibid. 16) were preferred by participants reading the ST and the fully post-edited text, while the text produced through translatability rules was preferred by the translators (who were tasked with evaluating quality and predicting PE effort) and assimilation users who were reading it in raw MT form. This indicated that the same rules that lead to a good user experience with the ST do not necessarily result in a good user experience with machine-translated texts.

In yet another group of comparison studies, the scope of comparisons of different modes of translation was expanded to include the evaluation of meta information on the speed and costs of translating in those different modes. This multifaceted evaluation, referred to as a *recipient evaluation*, was meant to "test the hypothesis that people will accept a lower quality of translation for some types of texts if the price is cheaper and the turnaround time faster", as stated by Bowker and Eghoetz (2007). The audiences that have been tested include university personnel (ibid.), speakers of minority languages (Bowker 2009), Spanish-speaking newcomers to a Canadian city (Bowker and Buitrago Ciro 2015), and Swiss Post personnel (Girletti et al. 2019). Equipped with knowledge of the cost and speed required to produce different modes of translation, the studies display a general tendency to prefer post-editing over HT or raw MT.

Two important factors make the scope of the studies described in this section different than the scope of this dissertation. First, for the most part, these studies focus on *potential* users of MT, while this dissertation is interested in how people *actually* use MT in real life. Second, the evaluations and experiments in these studies involved specific texts. Some texts were shorter, for example, individual sentences (Doherty and O'Brien 2009), others were longer texts (e.g., Fuji et al. 2001; Gaspari 2006). Some were in a slightly different form such as a MOOC (Hu, O'Brien, and Kenny 2019). Nevertheless, the experiments always involved a specific text. Conversely, this dissertation explores the practice of using raw MT and people's

perception of the practice, but not users' reactions to and reception of specific machine-translated texts.

3.2.3 Qualitative studies

The first study on users of raw MT, Henisz-Dostert's (1979) study on raw MT use by scientists from 1965–1973, unfortunately remains one of the few qualitative studies on people who employ MT for assimilation. Henisz-Dostert's study was one of the most important sources of background information for this dissertation for several reasons. First, it covered a broad array of questions on the informants' experience, including questions on reading and understanding MT; perceptions of usefulness, speed, and linguistic quality; and even a few questions on how informants experienced cognitive processes, such as "If the style of the MT is awkward, can you correct it mentally?" and "Do you get 'used to' reading MT?" (ibid., 193). No research since Henisz-Dostert has covered such a wide spectrum of the use of raw MT. Second, this study was important because, as mentioned above, there have not been many qualitative studies on raw MT use. A final reason was that, when I started the dissertation, it was the only in-depth study of a specific professional group of MT users that had been performed. It was tremendously helpful in the earlier stages of my work, but even more so when I was planning and carrying out the study that led to Articles II and III, which was a similar in-depth study of a group of professionals who employed raw MT.

Three articles from the early days of consumer-targeted MT included a qualitative evaluation of user reactions in addition to their descriptions of technical MT solutions. Flanagan's (1996) description of the different use cases covered by CompuServe's system of the time also included a description of the types of reactions they had received from system users, mostly in e-mails. Based on this data, Flanagan described a typical learning curve in using MT, from amazement to dismay to reconsideration to a final pragmatism upon learning what the system is capable of and where it does not perform well (Flanagan 1996, 193). In order to speed up this process, Flanagan recommended that efforts be put into education on the system. Two studies (Yang and Lange 1998; 2003) reported on the online tool AltaVista Translation with Systran, later renamed to Yahoo! Babel Fish. Similarly to Flanagan, these articles both described the technical solution and reported on user feedback, which in Yang and Lange's articles consisted of 5,005 e-mails collected in the spring of 1998. Contrary to Flanagan, the articles by Yang and Lange described users as

having realistic expectations for the service. These articles also offered an important contribution to the field. Based on their analysis of the feedback, they constructed the first categorization of types of MT usage: for assimilation, dissemination, communication, entertainment, and learning (Yang and Lange 1998, 281–282).

Most recently, the Health in my Language (HimL) project conducted a qualitative study of Polish and Romanian speakers in Scotland before starting to implement an MT solution for translating health information into those targeted languages (Birch et al. 2018). Interviews were held with potential users of the system to better understand their information needs and views to receiving information via MT. This study also included an evaluation of MT output by informants. The study concluded that the errors made by MT made raw MT unacceptable in the eyes of the potential users.

3.3 MT gisting in professional contexts

Besides the use case of individual companies employing raw MT for specific internal documents (Morland 2002) or for free use by employees (Smith 2003; Nuutila 2005; see Section 3.2.1 for more on the aforementioned), MT for gisting has been evaluated, piloted, or adopted more widely in various industries. In the area of intellectual property rights (IPR), the use of raw MT is a long-standing practice and is supported by industry-specific technology as well as official guidelines. Another area, customer support, has seen a general tendency to adopt the practice of publishing certain information via raw MT. A variety of fields have engaged in testing and piloting raw MT, including medical, government and legal. At least one field, education, shows evidence of wide MT adoption (for example, see Section 5.1.2) without any official support or provision of MT systems for employees and students. Although Church and Hovy (1993, 256–257) lamented that "it is very hard to find very many other natural-occurring limited domains that people care about, and consequently, this strategy [using MT in niche applications] is unlikely to be repeated very many times in the future," MT gisting in these professional contexts is in some cases widespread and has endured for years. Despite this long-term practice, the body of research on raw MT users remains limited.

3.3.1 Raw MT and intellectual property rights

The IPR field has relied on raw MT to access patent documents in languages that are not known by patent professionals for many years. In fact, this might be the professional group whose use of MT has been the most diffuse and long-lasting, surpassing even translators' use of MT for post-editing. A survey of translation professionals conducted in 2012 revealed that only 42% of respondents used MT in their work (Gaspari, Almaghout and Doherty 2015, 346). In an article in the *World Patent Information* journal that same year, List stated that "I'm sure everyone who is reading patents from China, Japan and Korea are reading them via MT" (List 2012, 193), suggesting that a very large majority of patent professionals were already using MT gisting on a regular basis in that same time frame. Despite the longevity of use, however, little attention has been paid to this use case outside of the IPR industry and research community.

Within research on IPR and patenting, the use of raw MT has sometimes been discussed in articles whose main focus is another aspect of patenting. Some articles that described technical solutions for patent MT also included information on aspects of patent professionals' work and their MT gisting, for example, Tinsley et al. (2012) and Rossi and Wiggins (2013). Similarly, a few survey studies of various aspects of patent professionals' work also included questions on their use of MT (Joho et al. 2010; McDonald-Maier 2009). A short description of the ways patent searchers access MT was provided by Tinsley (2017), who also discussed some of the challenges in developing MT for patents. One experiment with machine-translation information was also conducted by Tinsley et al. (2012). In the experiment, patent professionals were asked to evaluate machine-translated patent texts for quality, but also for relevance to a given case. Since one of the most frequent decisions patent professionals make concerns the relevance of documents, this was a viable test of whether MT met informants' needs and sufficiently disclosed the information in the patent document. The experiment found that in 73% of cases, patent searchers correctly identified relevant documents (ibid., 6).

3.3.2 Raw MT use in customer support

One of the early applications that large corporations found for raw MT was in customer support scenarios. Companies such as Intel, Microsoft, Autodesk, Symantec, IBM, Adobe and Cisco offered the information in their knowledge databases, support websites, and community forums to customers who preferred to

read information in their own languages through raw MT, as reported in both industrial newsletters and academic research. Several studies described solutions implemented by individual companies (Burgett and Chang 2008; Flournoy and Rueppel 2010; Thicke 2011; Wendt and Garcea 2013). A few articles in industry magazines (Dillinger and Gerber 2009; Thicke 2013) as well as at least one recorded conference panel (Burgett et al. 2012) covered a wider spectrum of activities in different companies.

A second way companies have implemented MT gisting in customer support is in direct customer assistance through e-mail and chat (e.g., Burgett et al. 2012). Eventually the customer support use case for raw MT led to productification and at least one company, Unbabel[11], that focuses specifically on offering solutions for customer support based on MT and rapid post-editing.

Various types of user studies have also been conducted, including general surveys (Burgett 2015, see Section 3.2.1) as well as end user evaluations of machine-translated material (Burgett and Chang 2008; Stewart et al. 2010; Thicke 2011; Thicke 2013; Wendt 2012). As reported by Thicke (2013, 50), companies quickly learned not to ask about linguistic quality, but rather "Did this translation answer your question?" Burgett and Chang (2008, 18) reported on such a survey at Intel. 44% of respondents who read raw MT reported that their question was answered, as compared to 47% of respondents reading human translation and 56% of those who were reading information in the source language of English.

3.3.3 Emerging areas for MT gisting: Health care and legal

A few emerging areas in which MT for gisting is being tested or piloted are the health care and legal fields. I use *health care* here as an umbrella term for a variety of areas including medicine, public health, and life sciences. Three recent literature review articles of the area gave overviews of available research. Two, Dew et al. (2018) and Vieira, O'Hagan and O'Sullivan (2020) involved systematic reviews of studies on various areas of health care. The third, Liu and Watts's (2019) article on mobile MT, included a section reviewing studies on the use of mobile MT in health care.

The current situation with MT in health care could be described as being in a state of testing, as there are few cases of documented or studied real-life use of MT in the context. Dew et al. (2018) reported that, of the 27 articles reviewed in their

[11] Unbabel's website: https://unbabel.com/

research, only 1 focused on a system that was in use, a solution that gathers and delivers information on global public health issues (Blench 2008). All others involved empirical tests or pilots of MT systems, prompting Dew and colleagues to state that "Overall, there is an immense need for real-world deployments and validation studies" (Dew et al. 2018, 65). They concluded that this field's need for accuracy in MT output is hindering the further adoption of MT, a viewpoint also reflected in other studies. The contexts in which experiments have been conducted include emergency response (Turner et al. 2019), health communications (Das et al. 2019; Khoong et al. 2019; Birch et al. 2017, 2018; Blench 2008), and doctor–patient or nurse–patient communications (Oladosu and Emuoyibofarhe 2012; Albrecht et al. 2013; Villalobos et al. 2017; Bouillon et al. 2017; Moberly 2018a, 2018b). A common theme in the results of the studies was caution because of the risk of mistranslation. In addition to research on MT, several studies have been conducted evaluating other technology or devices aimed at facilitating multilingual communication between doctors or nurses and patients, or at providing health information (see Turner et al. 2019; Parra et al. 2018; Bouillon et al. 2017; Villalobos et al. 2017; Albrecht et al. 2013). These solutions involved text, speech, or sometimes both, and relied mostly on pre-created phrases instead of translation.

Although academic research currently shows little evidence of MT being actively used in health care, commercial channels are reporting on areas where MT gisting is in use. One example of this is pharmacovigilence, defined as "the process of detection, assessment, understanding, and prevention of any adverse events or drug-related problems" (Iconic 2020, 5). In that process, reports from the side effects of pharmaceutical products are collected globally, resulting in a considerable amount of information in multiple languages. Raw MT is used automate the initial translation of those reports into English so that they can be categorized for further processing, as reported by Tinsley in the Globally Speaking Radio[12] podcast (Beninatto and Stevens 2019, 9).

The legal sphere has recently seen numerous cases in which the use of MT gisting during legal processes has resulted in court cases (see Vieira, O'Hagan and O'Sullivan 2020; Scott and O'Shea 2021). However, there are also legal processes in which raw MT has found its place. The first, patenting, is discussed in Section 3.3.1. The second use case is an emerging one: discovery. This is "a pre-trial phase in legal proceedings that involves the discovery and exchange of evidence and legal information between the parties" (Vieira, O'Hagan and O'Sullivan 2020, 8). Finding the relevant evidence requires reviewing very large amounts of information in multiple languages to

[12] Globally Speaking Radio podcast: https://www.globallyspeakingradio.com/

identify the information that is most relevant to the case being discussed, making it similar to the review of patent documentation that is required in IPR processes. Another similarity in these processes is that relevant information might be sent for human translation when necessary (ibid., 9), but raw MT is starting to be seen as acceptable up to a certain point even in legal settings (Tinsley in Beninatto and Stevens 2019, 9).

A final observation from this review of raw MT use in the health care and legal fields is that it does not seem easy to classify entire industries as suitable or not suitable for MT gisting. Based on the cases discussed here, the areas in which MT gisting is being employed are in very specific processes or sub-processes. This points to a clear need for granular analyses of processes to identify ongoing and potential use cases, supporting Church and Hovy's (1993) claim that MT is best used in niche applications.

Overall, focus on MT gisting in the workplace is a welcome contribution to the literature and our understanding of how people use raw MT. Studies on MT use in specific contexts, such as the IPR field discussed here and in Articles II and III of this dissertation, can help us to anticipate and prepare for emerging use cases in other contexts.

3.4 Characteristics of "good" use cases for MT gisting

Both academic and industry publications have examined or proposed what characteristics define a successful application of raw MT. The earliest article that considered the characteristics of appropriate use cases was Church and Hovy's 1993 article "Good Applications for Crummy Translation." In the article, Church and Hovy first reviewed several proposed applications of MT, then outlined six "desiderata" for a good niche application of MT (ibid., 256):

- it should set reasonable expectations,
- it should make sense economically,
- it should be attractive to the intended users,
- it should exploit the strengths of the machine and not compete with the strengths of the human,
- it should be clear to the users what the system can and cannot do, and
- it should encourage the field to move forward toward a sensible long-term goal.

Two studies in the early 2000s examined use cases they considered to be successful and then analyzed what characteristics contributed to that success. MT. Flournoy and Callison-Burch (2000) discussed factors in the MT-based Amikai system that contributed to its success: users' expectations were modified through education; the system focused on a narrow domain; it was intended for MT-mediated communication via chat, "a forum in which content is cheaply produced and has a very short shelf-life" (ibid., 2); the system had built-in functions for educating users; and finally, it also included a "Huh?" button with which a user could indicate to their fellow conversant that they did not understand a specific message. The second study that discussed aspects of good applications for raw MT was Ogura et al. (2004), which suggested that, although successful MT-mediated communication was mostly reliant on quality MT output, other factors were also important. They offered three examples of important factors: the tasks or goals of conversants, the subjects to be discussed using MT-mediated communication, and the "profile of the community" of users (ibid., 601).

Way's 2013 article "Traditional and Emerging Use-Cases for Machine Translation" reviewed ongoing and emerging use cases for MT gisting, then outlined parameters for evaluating whether a particular text would be best translated through human translation, post-edited MT, or raw MT (Way 2013, 2; also included in Way 2018, 160):

> In our view, the degree of human involvement required – or warranted – in a particular translation scenario will depend on the purpose, value and shelf-life of the content. More specifically, we assert that in all cases, the degree of post-editing or human input should be clearly correlated with the content lifespan.

Of the three parameters offered for evaluating the appropriateness of a text for human or machine translation, one (the shelf-life of the content) is concrete while the other two are vague. *Value* could refer to financial value or the value of the content in a reader's processes or life. *Purpose* could also refer to a variety of things, including the reason the content should be translated, the purpose a reader has for wanting to read it, or what the reader might do with the content (read it, save it for later use, etc.). The vagueness of the concepts makes them difficult to apply.

Instead of proposing what the features of good applications of MT might be, Castilho and O'Brien (2016) examined the factors companies reported actually using when profiling content. They conducted a survey of six multinational IT companies and asked if and how they profiled their content and how they assigned the mode of translation required for each. Asked what factors were involved in decisions on if and how different content types would be translated in their companies, participants'

answers centered on the factors of brand image, the business case at hand (cost, profit, return on investment, revenue), geography (region, size and growth of market), user-centric factors (behavior, audience, rating), and product-specific factors such as the volume of the content to be translated and the effort involved in translating it (Castilho and O'Brien 2016, 28). The article noted that companies reported using raw MT solely with technical information, and that two factors appeared to be prominent in the decision to use raw MT: users' expectations for quality were not high and the content type in question had not been previously translated.

In their MT Success Guide, Iconic Translation Machines proposed a model for evaluating potential use cases for MT (Iconic 2016), which lists eight concrete factors the company applies when assessing the suitability of translation projects for MT. Seven of these factors concern technology and the business case at hand, including the languages required, the volume and type of content to be translated, the level of quality required, the type of data available for training the MT engine, the effectiveness of any currently used TM systems, and the technical environment the MT system would need to be integrated into. The final factor is contextual: the experience of the party buying the MT solution, which Iconic suggests "arguably has the single biggest impact on successful (painless) adoption of machine translation"[13]. When the case involves raw MT, the experienced buyer has a better understanding of how to evaluate when MT might be fit-for-purpose in their own contexts. In addition, Iconic claims that buyers with less experience with MT tend to be more sensitive to errors in the MT output.

In their model on evaluating the type of translation that would be appropriate for specific content types, Nitzke, Hansen-Schirra and Canfora (2019) included considerations of the risk and sensitivity of text types, the quality levels produced by the MT solution to be used[14], the time and manpower available for translation, the lifespan of the texts to be translated, and the application of controlled language in text production. Another factor they included was the size of the audience for the translation, under the logic that light PE or raw MT might be suitable for a text with a small audience, but a large readership requires higher quality.

The evaluation of use cases for MT was also taken up by the Translation Automation User Society (TAUS), which proposed six considerations for evaluating potential applications for MT (Panic 2020). Similar to the Iconic model, this model

[13] Iconic blog: https://iconictranslation.com/2016/08/mt-success-series-5-buyer-experience/

[14] Includes considerations also mentioned by Iconic (2016) and TAUS (n.d.) such as required languages and available training data.

contains technical considerations (the language pairs involved and linguistic quality of MT output), business considerations (the use case or business scenario involved and considerations of costs and return on investment), and considerations of the suitability of the content for MT. They focused on three factors to assess suitability of content: *utility*, or the function of the content; *time*, referring to how quickly translation is needed; and *sentiment*, referring to whether content is "meant to emotionally engage the reader about a company or product" (ibid.).

Bowker (2020) analyzed the idea of "fit-for-purpose" translation, which she considered to be a way to define the quality level required in a translation project. Bowker listed the factors that should be considered in such an evaluation, taking the factor of *perishability of content* from Way (2013) and the factors of *audience* and *purpose* from Durban and Melby (2008). Unfortunately, the latter concepts remain vague. Neither Bowker nor Durban and Melby discuss what specific aspects of the audience might affect decisions on how to translate content. The explanation of *purpose* is given as: "Sometimes all you want is to get (or give) the general idea of a document (rough translation); in other cases, a polished text is essential" (Durban and Melby 2008, 3). While this covers one aspect of purpose – the desired level of understanding for the reader – a number of other aspects of *purpose* might also affect translation, for example, the motivation for having a text translated, what exactly will be done with the translated text (own use, share with someone else, save), and which process the translated text will be used in.

In reviewing the various proposed characteristics of an appropriate use case for raw MT, certain themes emerge. The first is the user. Researchers have discussed what types of audience will buy or read a text (Iconic 2016; Durban and Melby 2008; Bowker 2020), the expectations of those audiences (Church and Hovy 1993; Flournoy and Callison-Burch 2000), the size of the audience (Nitzke, Hansen-Schirra and Canfora 2019), and the attractiveness of a translation system to users (Church and Hovy 1993). In regards to the source text, researchers have discussed its perishability (Flournoy and Callison-Burch 2000; Way 2013 and 2018; Bowker 2020), whether it was produced using a controlled language (Nitzke, Hansen-Schirra and Canfora 2019), and the sensitivity and risk levels of the text (Nitzke, Hansen-Schirra and Canfora 2019). Finally, more abstract factors such as purpose (Way 2013; Bowker 2020) and value (Way 2013) have been proposed as parameters in considering whether MT would be suitable or not, but without a detailed discussion of exactly what aspects of these concepts might affect the use of MT.

One aim of this dissertation was to identify factors that influence the use and reception of raw MT. Since such factors also act as good signposts for helping us

define characteristics of good use cases for MT gisting, they overlap somewhat with the types of characterics identified here.

3.5 MT gisting and risk

The risks involved in human translation has received some attention in Translation Studies literature. More recently, focus has turned to MT and the risks involved with using it for both post-editing and gisting.

Research on risk and human translation has focused on both the individual translator and on the full translation process and the various actors involved in it. Both Pym (2015) and Pym and Matsushita (2018) examined the risk faced by individual translators. Pym (2015) identified risks concerning the translator's credibility, which involves readers' trust in them; risks associated with uncertainty when making translation decisions; and the communicative risk of non-cooperation, which can occur if communication fails. Pym also pointed out that risk management is not only about negative consequences, it also includes the evaluation of the benefits that can be gained through communication (ibid., 71). Pym and Matsushita (2018) analyzed the various ways a translator can mitigate the risks inherent in their processes. Expanding the focus to encompass the full translation process, Zaveckaite and Ulbinaite (2018) examined risk management in language service provider companies and promoted the employment of risk management processes, while Canfora and Ottmann (2019) introduced their model of Risk Management for Translations, which they claim accounts for all agents in the translation process, including end users of translations.

Some of the research on risk and MT has similarly applied risk management principles to MT production and use. Two studies endeavored to both examine risks associated with MT and outline methods for mitigating those risks. Canfora and Ottmann (2020) examined risks associated with the use of MT by translators as well as by non-translator employees in companies. The types of risks they uncovered included translation errors, questions of liability, and data security risks tied to the use of free online translation tools by both translators and others who might put company-sensitive information into such tools. They suggested methods for countering these risks through both non-technical (e.g. legal agreements and laws) and technical measures. They also stressed the importance of using risk management to mitigate risks. Martindale (2020), focusing more specifically on MT gisting, similarly identified risks concerning errors in MT output and then outlined mitigation

strategies for addressing those risks. Martindale identified the following risks: MT output is believable, there is a lack of means or motivation to verify information, and the raw MT will be used to inform action (Martindale 2020, 26). Like Canfora and Ottmann (2020), Martindale also categorized mitigation strategies into technical interventions and non-technical policy interventions. Risk management on a personal level has also been addressed. Bowker and Buitrago Ciro (2019, 88) argued that an MT user with high MT literacy would "appreciate the wider implications associated with the use of machine translation," implying that they would evaluate the risk involved in using raw MT and take necessary precautions.

Nitzke, Hansen-Schirra and Canfora (2019) employed risk management in a decision-making model focused on evaluating whether a particular text or text type would be best translated in a fully human process, with MT and post-editing, or if it could be handled by raw MT. As mentioned earlier (Section 3.4), the model included various elements, including some elements of risk, for example, the sensitivity of the MT training data and the risk level of the text to be translated. However, the first question in the decision tree is "Can I use MT? Do the benefits of MT+PE outnumber the risks?" (ibid., 246). While it is a valid question, the model does not provide tools for evaluating this question, and it seems it should be answered before advancing in the decision-making. The role of that question in the model is not clear.

Another angle of risk that has been studied are the risks involved in people trusting MT output too much. Martindale and Carpuat (2018) found that fluency in MT output increased informants' trust in MT output even when the output contained errors. Rossetti, O'Brien and Cadwell (2020) tested users' comprehension of and trust in machine-translated messages related to a weather crisis and had similar results to Martindale and Carpuat: better comprehension of machine-translated texts tended to be associated with a higher level of trust in the texts. These results are important especially in relation to risk and NMT, which is known to produce output that, although fluent, may contain errors (see Castilho et al. 2017).

Finally, one group of studies has addressed the risks associated with using MT in contexts that are viewed as highly sensitive. Scott and O'Shea (2021) highlighted the risks of using MT in legal processes. Vieira, O'Hagan and O'Sullivan (2020) explored the risk involved with using MT in the medical and legal fields. Through investigating research literature from those areas, and not research in MT or Translation Studies, they found evidence that MT was being adopted, but that the understanding of the risk involved with it was low. Canfora and Ottmann (2020) similarly highlighted that in safety-critical domains, raw MT use is so risky that it should be avoided.

The inherent inevitability that some MT output will be erroneous (see Castilho et al. 2017; Koponen and Salmi 2015) dictates that risk is a constant companion for users of raw MT. It is likewise important for researchers to focus on how users conceptualize and consider that risk.

3.6 MT gisting and distributed cognition

In Article III of this dissertation, I applied the principles of distributed cognition to analyze patent professionals' use of raw MT (Nurminen 2020). As described by Muñoz Martín (2010, 152), distributed cognition refers to situations in which "meaning emerg[es] or [is] built from the interaction of various mental processes and the environment." Clark and Chalmers (1998) employed the term *extended cognition* to describe how a human brain and an external entity can be combined to form a cognitive system, in which actions and behavior are governed by the system and, "If we remove the external component the system's behavioural competence will drop, just as it would if we removed part of its brain" (ibid., 8–9).

Aside from a few questions related to cognition in the earliest study on MT assimilation users (Henisz-Dostert 1979; see Section 3.2.3), the body of research on raw MT reception reveals little focus on distributed cognition or other aspects of cognition. O'Brien's (2017) chapter on MT and cognition in the *Handbook of Translation and Cognition* is devoted mostly to cognition in the post-editing process; very little mention is made of cognition as it relates to users of raw MT. O'Brien states clearly that this is due to the research in MT gisting being in its early stages (ibid., 313). Raw MT users are mentioned briefly in a section on evaluation methods, in which O'Brien (ibid., 327–328) warns that

> [t]he evaluation methods discussed in Section 17.2 are taken, to some extent at least, as acting as a proxy for measuring the usability and acceptability of MT output by those for whom it is essentially produced: end users. However, these are not direct measures of interaction with MT and tell us very little about the cognitive processing effort that might be involved in reading, comprehending, and acting on a set of machine-translated task instructions.

One area of MT user research has examined aspects of cognition in the reception of machine-translated texts, either raw or post-edited, through eye-tracking. Doherty and O'Brien (2009) first tested how well eye-tracking works as a method for evaluating MT output, hypothesizing that poor output will result in a higher cognitive load, which will show in eye-tracking metrics. They found that eye tracking

correlated "reasonably well" with human evaluation (Doherty and O'Brien 2009, 214). Since then other studies (Doherty 2012; Doherty and O'Brien 2014; Castilho et al. 2014; Castilho 2016; and Hu, O'Brien and Kenny 2017) have employed eye tracking. Nevertheless, in these studies eye tracking and the resulting measure of cognitive effort were employed to evaluate translated content, and the focus was on comparing the output of different translation methods (human, post-edited, or raw MT) rather than to explore aspects of cognition.

A small group of studies explored or applied the method of gap filling as a way to measure "the usefulness [of machine-produced text] in understanding the meaning of the original text" (Ageeva et al. 2015, 137). Somers and Wild (2000) tested the method for evaluating the output of different MT systems, concluding that the method was a reasonable method for measuring readability. Ageeva et al. (2015) and Forcada et al. (2018) compared gap filling against the more traditional questionnaires to measure comprehensibility. The results indicated that gap filling led to results which were reasonably similar to those attained using questionnaires, but Forcada et al. (2018) points out that gap filling is a cheaper option.

A separate but related area of research that provided important background material for this dissertation was research on cognition in human translation processes. Because research on cognition and raw MT reception is scarce, I needed to rely on work from related areas for theoretical insights. Theoretical work on translators' process proved useful for this purpose. Especially work in the area of situated and distributed cognition seemed applicable to the case of MT gisting. Risku (2014, 335) discussed "situated, embodied and extended cognition" in relation to translators' work and found that "cognitive processes are context-dependent, i.e., they are dependent on and partly constituted by the social and physical environment in which they are carried out." Distributed cognition in relation to the translation process was described by Muñoz Martín (2017, 564) as follows: "Meaning is encyclopedic, and it is a process, not a thing…Understanding is an activity that crucially depends on the environment – and also on experience – because environmental affordances foster and constrain meaning construal."

A final contribution that research on cognition in translation processes has made, both to this dissertation and Translation Studies research, was a reinforced emphasis on qualitative and workplace-based research. Risku (2014, 335) argued for more diverse methodologies, stating that besides experimental study, "we also need other methodological paths of inquiry to model the cognitive processes in translation and to establish a deeper understanding of how translations are produced." Ehrensberger-Dow (2019, 41) also emphasized that study in this area should rely on

both experiments in laboratory settings and workplace observations. Indeed, the qualitative, workplace-based research called for by Risku and Ehrensberger-Dow, and implemented in Articles II and III of this dissertation, was effective in leading to a better understanding of the cognition that takes place in raw MT reception.

3.7 MT and accessibility to information

One of the prerequisites for full participation in society is access to important information, such as information on education, health, and governmental services. In fact, the right to information can be seen as one of the fundamental human rights derived from agreements such as the United Nations Declaration on Human Rights (United Nations 1948) and the European Convention on Human Rights (Council of Europe 1950). However, in addition to being available, information also needs to be accessible, including being linguistically accessible to people who do not speak the languages the society uses for publishing information. The groups that are affected by such language barriers might include newcomers to a society, refugees, and people in crisis situations. These groups could be provided with access to important information through translation which, as noted by Suojanen, Koskinen and Tuominen (2015, 57) "is often an accessibility service, allowing readers access to a text which they might not have been otherwise able to understand." However, the resources required to translate into multiple languages quickly become prohibitive.

The idea of using MT as a partial answer to these issues was first discussed in the early years of MT development. Hutchins (1986, 15) described researchers being motivated by ideas of removing language barriers, delivering information to developing countries, and promoting international cooperation and peace. The topic was approached again later, when the need for MT solutions for minority languages and immigrant populations was recognized (see Carbonell et al. 2006). The line of research on these "low-resource" languages, which do not enjoy large amounts of the types of linguistic resources needed to build MT solutions, continues today. For example, two workshops devoted to low-resource languages have been held in conferences hosted by branches of the International Association for Machine Translation (Liu and Karakanta 2018; Liu 2019).

Another important line of research concerns accessibility to information in crisis scenarios. The International Network on Crisis Translation (INTERACT) project[15] initiated the conversation on the role of translation in crisis scenarios, including the

[15] INTERACT project: https://sites.google.com/view/crisistranslation/publications

role language technology such as MT can play. Various aspects of the question have been investigated, including ethical considerations in the use of MT and language technology for crisis communication (Parra Escartín and Moniz 2020; O'Mathúna et al. 2019). Today, crisis translation and MT is a growing field of research (see Federici and O'Brien 2020).

The growing discussion on ethics in MT (see Kenny, Moorkens and do Carmo 2020) acted as a catalyst for this dissertation to consider ways in which MT could be used for improving people's lives (see Article V). It is important to continue to promote the idea that, alongside its other uses, we should also strive to find ways to use technology for the good of humanity, and that this quest is an integral component of ethics.

3.8 Work in related areas

In addition to the literature discussed thus far, which focused specifically on MT users and reception, this dissertation also relied on limited literature from the areas of internet studies and user experience. The influence of these fields on the dissertation is greater than what might be immediately discernible.

Especially in the initial phases of the dissertation, internet studies provided me with inspiration, assurance that my ideas for research were valid and could be implemented, and guidance for defining the types of goals, research questions and designs, and methods that I thought could be applied to users of raw MT. When I began to explore the literature on MT and usability, I quickly noticed that the type of research I was considering was not well represented. For example, my forays into usability literature repeatedly resulted in laboratory-based studies on users' reactions to specific products or tools, which was not what I wanted to do. Internet studies, on the other hand, reflected my own aims. In fact, the context in which internet studies was born mirrored the situation in MT when I began my research. The focus in early internet research was mostly on technology until a different type of researcher began to ask new questions. This new approach was described by Ess and Consalvo (2011, 1) as follows: "we and our colleagues seek to study the distinctive sorts of human communication and interaction facilitated by the Internet." Similarly, the bulk of the research in MT when I started my research was devoted to the technologies that are behind MT, with a new but still small body of studies on how one group, translators, used the technology. There was little knowledge on other MT users or the broader context of the human communication processes that they were

using MT in. In other words, and borrowing another concept from internet studies, there was a lack of focus on how MT was used in "everyday life" (Bakardjieva 2011). The literature gave me confidence that I should continue with my initial ideas.

A second influence of internet studies was in approach and method selection. It informed the choice of inductive, qualitative methods for the study on patent professionals (Articles II and III) as well as the methods employed in Article IV. That article piloted the gathering of data using MT-mediated interviewing, which involved conducting interviews via instant messaging (IM) with integrated MT. While material on that specific method was not available because it was being piloted for the first time, internet studies research on interviewing via IM proved helpful. Research by Markham (2004) and Voida et al. (2004) explored the nature, advantages and disadvantages of interviewing via IM, while other studies (for example, Kazmer and Xie 2008; Opdenakker 2006) compared IM interviewing with other common interviewing methods such as face-to-face, e-mail, and telephone interviewing. The challenges associated with the method have also been discussed. For example, researchers are accustomed to relying not only on speech but on other social cues during face-to-face interviews. The lack of such cues in IM interviewing was noted as being potentially detrimental (Oppdenakker 2006, Markham 2004; Voida et al. 2004).

The most visible influence of user experience studies on this dissertation was the contribution of the concept of context of use as defined in ISO standard 9241-210:2019 (ISO 2019) and described in Section 2.2. Context of use was also discussed in Suojanen, Koskinen and Tuominen (2015), which applied the principles of user experience and the related concepts of usability and user-centered design to translation, developing the concept of user-centered translation (UCT). Similarly to internet studies, the field also influenced my work in ways that are not as noticeable. Its focus on the users of technology and translation provided me with a consistent encouragement.

4 RESEARCH APPROACH, PROCESS, AND METHODS

4.1 Overall approach

When approaching an under-researched area such as MT gisting, one choice would be to work bottom-up, conducting studies on individual phenomena that would collectively eventually build a broader picture of the topic. A second choice would be to work top-down, first attempting to capture a broader picture of a topic and then identifying individual phenomena that could be studied further. This dissertation is of the latter variety; the intent was to first explore a variety of different MT gisting contexts and, through that exploration, identify individual phenomena that could be studied more deeply in future research. My aim was therefore to generate new questions and ideas for further study rather than to confirm existing hypotheses.

Two issues made it difficult to select the contexts and groups to study. The first was that the user group I wanted to study was extremely large and the second, that this group was significantly under-researched. On the one hand, since so little was known about who was using raw MT and how they were using it, gathering as much information as possible from as large a group as possible would seem to be a good starting point. On the other hand, in order to know which questions to ask such a large group, one would need to begin with a deeper understanding of how people are using raw MT, which can be attained through qualitative studies of smaller groups. A third difficulty, the question of locating and accessing specific groups that use raw MT, eventually led to the resolution of these issues. Simply put, in selecting the groups and contexts to study, I sought and used all opportunities that I could find. Depending on the nature of each, I then selected appropriate methods that would lead me in exploring my research questions.

The resulting collection of articles examined these groups using a variety of methods, many of which were inductive and qualitative. Only one of the studies was quantitative. The data-gathering methods included online surveys; interviewing via chat, Skype and in person; and convenience sampling and artifact searching. Data analyses were conducted using thematic analysis, qualitative methods, qualitative

meta-analysis, and quantitative methods. This chapter describes the progression of the dissertation work, the groups and artifacts studied, the data collected, and the methods applied throughout the dissertation project. It then discusses reflexivity and ethical issues.

4.2 Phases of the research process

The dissertation consists of five articles that were published from 2016 to 2020. They are numbered in this summary thematically rather than chronologically. Article I (Nurminen and Papula 2018) concerned a broad study of users of an online MT tool and provided a snapshot of use and users at a specific point in time. The subsequent articles then examined various specific user groups and contexts. Articles II (Nurminen 2019) and III (Nurminen 2020) provided a deep study of a group of professionals who use raw MT at work. Article IV (Nurminen 2016) examined MT-mediated communication between an interviewer and four different users of the same online MT tool as in Article I. Finally, Article V (Nurminen and Koponen 2020) provided an analysis of situations in which people suffer from a lack of accessibility to information they need, which might be provided with the help of MT.

Articles I–III are the most similar. Each of them examined real users of raw MT and aimed to explore and analyze participants' MT use. The research methods consisted of surveying and interviewing, both of which involved asking people to report on their own MT activities.

Articles IV and V are somewhat different than the first three articles. Article IV differed from the first three articles in three respects. First, while the people studied were actual users of raw MT (and also from the same user group as in Article I), the focus of Article IV was not on their reports of their MT gisting. Instead, the study involved interviews that were conducted via MT-mediated communication and it focused on analyzing the communication in the interviews rather than the content of the communication. A second aspect of the article that differed was that, as the interviewer in the study, I was one of the participants whose actions were analyzed. A third difference between Article IV and the first three articles was that Article IV concerned MT for communication, while the first three concentrated on MT for assimilation.

In fact, I did not decide to include Article IV in the dissertation until late in the project. I was finishing the article while concurrently applying for the Ph.D. program and its topic of MT for communication did not seem to fit into the scope of the

proposed Ph.D., which was MT for assimilation. I therefore declared Article IV out of scope and proceeded with research on MT for assimilation for the next two years. What eventually happened was that, while Article IV did not change over time, the scope of the dissertation did. As explained in Section 2.3, over the course of the project I realized that, rather than being considered two different phenomena, MT for assimilation and MT for communication are actually part of the same phenomenon of MT for gisting and should be considered as such, including in the dissertation. The dissertation's scope was therefore broadened and Article IV included.

Article V is also different than all other articles. First, unlike the other articles, it was not a direct study of raw MT use and users but rather a meta-analysis of articles and projects that focused on a specific group of current and potential MT users, namely, people who suffer from a lack of access to information they need. A second difference in Article V is that its scope included the use of post-edited MT as well as raw MT. A holisitic examination of the problem of accessibility and an evaluation of solutions that involved either raw or post-edited MT helped to highlight the factors that should be taken into consideration when deciding on the level of human intervention that is appropriate for different contexts.

4.3 Study subjects, data, and methods

Table 3. on the next page introduces the groups studied in the dissertation, the data analyzed, and the methods adopted in each article. Note that Articles II and III used the same data set but separate data analyses. The table is followed by a more specific description of the approach, data and methods involved in each article.

Table 3. Overview of study subjects, data and methods used in dissertation articles

Arti-cle	Group or artifacts studied	Description of group or artifacts studied	Data gathering methods	Data	Analysis method
I	Users of Multilizer's PDF Translator[16]	Large, diverse and geographically dispersed group of users of an online MT tool	Survey + compiling of log data	Survey answers (n=1579) + usage logs	Quantitative analysis
II	Patent professionals	Nine Scandinavian IPR professionals. One member of the group acted as the key informant	Semi-structured interviews at informants' workplaces or via Skype	Transcripts from interviews	Thematic analysis (qualitative)
III					Analysis of data through lens of distributed cognition (qualitative)
IV	Users of Multilizer's PDF Translator and myself	Four active users of an online MT tool and myself	Semi-structured chat-based interviews conducted via MT-mediated communication at workplaces	Communication with participants and own observations + transcripts from interviews	Qualitative analysis
V	Meta-analysis of existing research on groups who would benefit from greater access to information in their language	Existing research and information on projects that included well-defined use cases and target users	Convenience sampling and searching	Scientific literature, web sites, project documentation, personal communication	Meta-analysis

4.3.1 Article I

The research for both Article I and IV was conducted in cooperation with Multilizer, a Helsinki-based language technology company. Multilizer produces an online MT tool, PDF Translator, that machine-translates whole documents, and it has a large and geographically dispersed group of users.

The goal of Article I was to produce a snapshot of PDF Translator users in a precise, short period of time. It was meant to provide a view into MT use that was

[16] PDF Translator (currently Multilizer Document Translator): pdf.multilizer.com

more broad than deep. During a four-month period, we collected the logs from everyone who requested a translation in the tool and we also offered each requester the opportunity to answer a survey. With the goal of obtaining as many responses as possible, we devised a short survey that required an average of only three minutes to answer. It included four questions on demographics and four on informants' use of MT. Six language versions, reflecting the six most popular language variants of the tool, were available. The survey resulted in a total of 1,579 responses.

Results were analyzed by applying quantitative methods. We examined logs to uncover where people were using the tool and which languages (both source and target) were requested during the study period. We compiled the responses to each survey question individually, after which we analyzed the responses regarding MT in terms of different demographic variables. The resultant data allowed us to examine differences in MT use according to factors such as age and language.

The log information reflected a random sample of PDF Translator users, while the responses to the survey can be assumed to reflect users who have sufficient interest in the tool and MT to respond to a survey. During the study period, the tool was accessed by people in 181 countries and territories, and survey responses were received from users in 97 countries and territories. While this geographical distribution is exceedingly broad, it nevertheless constitutes only a sample of the users of this particular online translation tool for PDFs and is therefore not representative of the general populace.

4.3.2 Articles II and III

The study that led to Articles II and III focused on a small group of Scandinavian patent professionals who consult raw MT in specific processes at work, and this places the research in the realm of workplace studies. Although the study focused on the reception instead of the production of translation, it was similar to workplace research on the translation process, for example, that described by Risku, Rogl and Milosevic (2019, 3–4) and Ehrensberger-Dow (2019).

More specifically, the study might be described as an ecosystem study. A business ecosystem is defined by Investopedia[17] as "the network of organizations – including suppliers, distributors, customers, competitors, government agencies, and so on – involved in the delivery of a specific product or service through both competition

[17] Investopedia: https://www.investopedia.com/terms/b/business-ecosystem

and cooperation." The concept is similar to the production networks described by Abdallah and Koskinen (2007) in that both involve an interdependence between organizations, joint goals of delivering specific products or services, and the presence of both cooperation and competition. However, whereas production networks as defined by Abdallah and Koskinen emphasize the binding of participants into one larger economic unit (ibid., 674–675), the IPR ecosystem described by the informants of Articles II and III involved organizations that retained and protected their own goals while cooperating to the degree necessary to produce valid, strong patents.

A second defining characteristic of the study that led to Articles II and III was that it was qualitative. Data was collected by conducting semi-structured interviews. In keeping with the focus on MT use at work, six interviews were conducted at informants' workplaces, while three interviews were conducted via the audio function of Skype. After the first few interviews, a followup interview with the key informant was performed to verify the interviewer's understanding of key terminology and processes before continuing with the rest of the interviews. The interviews were recorded and subsequently transcribed by first putting them through an automated transcription tool and then editing the results. Although manual transcription by the researcher is often considered beneficial due to the "intimate familiarity with your data that doing your own transcribing affords," (Merriam and Tisdell 2016, 132), I discovered that the editing done after automated transcription required such meticulous care and effort that I did indeed become familiar with the data and was able to begin analyzing it while post-editing. It nevertheless required less time than a fully manual method and therefore was more efficient.

I applied thematic analysis as defined by Braun and Clarke (2006; 2013; 2017; n.d.) in analyzing the data for Article II. This approach to qualitative analysis was flexible, which was useful in such an exploratory study, and it was also an explicit method with ample support material available. During this first analysis, I maintained an explorative and inductive approach and followed what I found in the data. To enhance the internal validity or credibility of my analysis, I employed a "member check" (Merriam and Tisdell 2016, 246) in which several of the study's informants reviewed a summary of the results of my analysis to verify that the findings reflected their own experience.

In Article III, I began with the same raw, coded data from Article II, but performed a new analysis in which I analyzed the data through the lens of distributed cognition. When I first read about situated and distributed cognition, I recognized that it might be applicable to the reception of raw MT by patent professionals. I set

out to conduct a deeper analysis of whether this were valid and whether distributed cognition would be a suitable means of describing the phenomenon of MT gisting in this use case. The analysis made it apparent that it was indeed an appropriate and useful lens through which to view MT gisting.

In retrospect, the use of qualitative, explorative methods was an even more important decision than I had originally anticipated when I embarked on the study of patent professionals. The state of research on the users of raw MT at the time was described by O'Brien (2017, 327) as "still in its early stages." The body of research that did exist was oriented toward administering surveys and laboratory-based, experimental methods. However, as mentioned previously in this dissertation summary and as highlighted by Tuominen (2012), at the early stages of research on a particular group or phenomenon, it is not always clear what types of questions we should be asking. Considering the nascent stage of the research on raw MT users, qualitative inquiry was an effective way to learn enough to be able to define and carry out further studies. A second benefit of qualitative research, as discussed by Saldanha and O'Brien (2014, 37), is that it is effective for generating hypotheses on what occurs in a specific situation, which can then be used to investigate other situations. I consider this idea of generating hypotheses that can be the bases for future research to be one of the main contributions of not only this study, but also the dissertation.

An important reason for the selection of both the scope of an ecosystem and qualitative methods was that it introduced a previously under-explored viewpoint to our understanding of raw MT use, namely, its use by groups. This viewpoint is important not only for increasing our understanding of groups' use of MT, but also because it offers a nuanced insight into our overall understanding of MT gisting. Suojanen, Koskinen and Tuominen (2015, 26) stated that a potential problem with an over-emphasis on user experience is that "it emphasizes the individuals at the expense of shared views and interactive co-construction of experience." Articles I, IV and V focused on the experience of individuals with raw MT, but the study on patent professionals led to the discovery of aspects of reception that are not individual but communal, while also identifying factors in the surrounding environment that influence reception.

4.3.3 Article IV

Article IV was the second study that was conducted through collaboration with Multilizer. It focused on a small group of four Spanish-speaking users of the PDF Translator tool, who were recruited through a message embedded in the tool. These users, who responded to a call for participation in research that would involve the innovative use of an MT tool, could be assumed to be enthusiastic users of technology.

The inductive study that led to Article IV concerned piloting an original method for gathering data: interviewing informants through MT-mediated communication. In it, I conducted semi-structured interviews using the chat function of Skype Translator, which provided a chat interface backed by MT. I typed my messages in English, informants typed theirs in Spanish, and the tool translated back and forth between the languages, always showing both source and target texts. One benefit of this text-based approach was that no transcribing of the data was required; once an interview was complete, the transcription was ready. As one of the goals of the study was to pilot an original method, my own experience of the interviewing process was considered part of the data. Unfortunately, I did not keep a diary, which means that this data source was not as robust as it could have been (for more on me acting as both researcher and participant, see Section 4.6.3)

The qualitative analysis of the data focused on identifying factors that influenced the interviewing situation and the interviews themselves. Some of the results were derived through closely examining the full interview process, from the first contact with potential informants to the finalized transcripts. Others were uncovered through inspecting the transcripts themselves. Through these processes, I identified seven considerations for interviewing through MT-mediated communication, which are discussed in Section 5.3.

4.3.4 Article V

Article V was a qualitative meta-analysis of artifacts such as scientific papers, web sites and documentation from research projects. These artifacts discusssed a specific group of targeted users of raw MT, those who need information in order to participate in the societies they live in, but who are blocked from that information because they do not understand the languages it is published in. The article therefore addressed a question of ethics.

We collected data for the article from different areas. In addition to literature related to MT technical development and accessbility from the viewpoint of public authorities, we focused on projects aimed at developing MT solutions for increasing accessibility to information. We began with a convenience sampling of projects that we were already aware of and then expanded the scope through searching, aiming to identify projects with a scope that included not only MT system development, but also well-defined use cases and target users. However, the search was unsystematic and one of our recommendations for future research included a more comprehensive approach for identifying efforts to increase accessibility to information through MT.

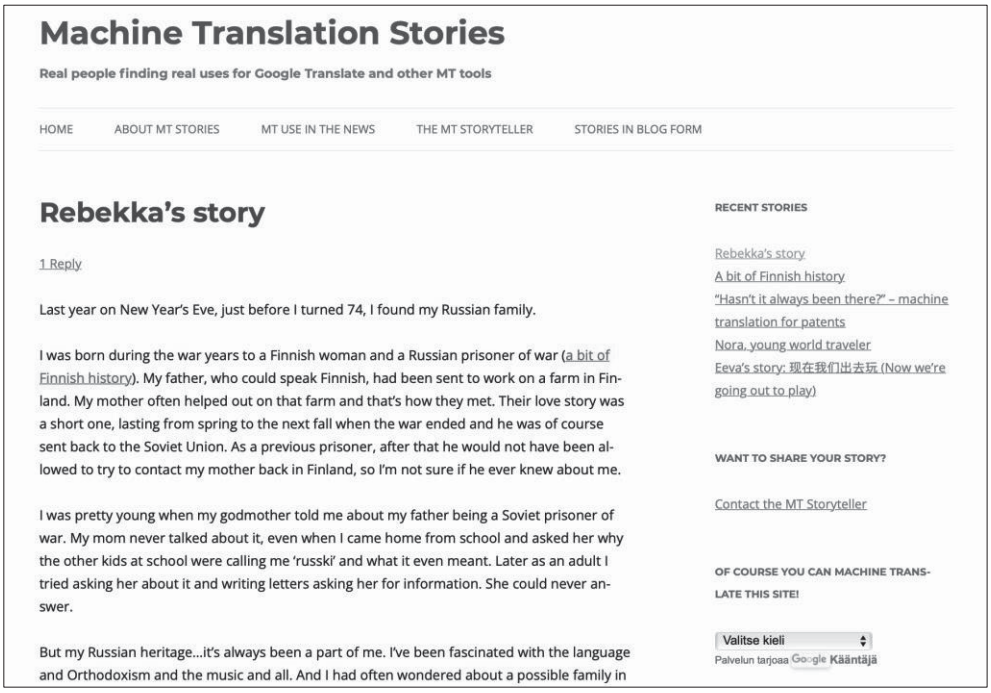
Qualitative meta-analysis is defined by Timulak as "an attempt to conduct a rigorous secondary qualitative analysis of primary qualitative findings" (2009, 591). In our case, we analyzed findings in a variety of academic articles and industry reports and derived themes related to the use of MT to alleviate accessibility issues. Our analysis first focused on the views of technical developers and public authorities on the need for increased accessibility to information and what their motivations and goals have been in introducing the idea of using MT for that purpose. We then analyzed literature on individual projects: how they were funded; what their goals and target groups were; whether they focused on MT for post-editing, MT for gisting, or both; what the results of any tests were; which solutions eventually resulted in implementation; and finally, any plans they had for future development.

4.4 Machine Translation Stories[18] as a method for idea and question development

Throughout the latter half of my dissertation work, I also worked on a side project that influenced the dissertation significantly. In the project, I developed Machine Translation Stories (MT Stories), a website that consists of small case studies of various individuals' use of MT, which are delivered as narratives. The website is depicted in Figure 3.

[18] MT Stories website: www.mt-stories.com

Figure 3. MT Stories website



A total of 16 stories were researched and written during the time of the dissertation. As the title indicates, they were written as stories, primarily as first-person narratives. The data was never analyzed as a set, nor are the stories meant to be read as academic texts. However, the data for most of the stories was gathered through a formal method of semi-structured interviewing followed by transcription, and the results would be appropriate for analysis and publication.

I had several motivations for developing MT Stories. First, I wanted to publish information on people's use of MT. When I began the dissertation, a very large number of people were already using MT, yet there seemed to be a gap in knowledge regarding the phenomenon: how people used it, for what purposes, in what contexts, etc. I realized that academic research takes time and I wanted to begin disseminating information more quickly. In addition, it would be beneficial to publish information in a different genre and medium than the eventual scientific articles.

A second motivation for creating the website was that I needed to begin formulating questions to ask when interviewing people for the dissertation research. When working in an emerging area of research, it is not always clear what types of questions might elicit the information and understanding desired. To address this, I

simply began speaking with people about it, even if I could not be sure that those discussions would lead to data that would be used in the dissertation. In retrospect, MT Stories proved instrumental as a sandbox for discovering new questions and experimenting with trial questions.

Apart from these premeditated reasons for the project, it made another, unexpected, contribution to my dissertation. The user group of patent professionals, which became the target of two of the dissertation articles, was originally discovered through interviews for MT Stories. When I heard from an acquaintance about this particular use of MT, I was prompted to search for someone willing to be interviewed for a story. After interviewing two such people, I realized that the user group was interesting enough to warrant further study and research. I subsequently embarked upon this study and it eventually led to Articles II and III. After both academic articles were published, I also published an MT story on a patent professional.[19]

4.5 Reflexivity

In qualitative studies, researchers are called upon to "articulate and clarify their assumptions, experiences, worldview, and theoretical orientation to the study" (Merriam and Tisdell 2016, 249). As they act as the main interpreters of the data, it is important to understand these aspects as part of the research process. This dissertation relied heavily, albeit not exclusively, on qualitative methods and therefore, the inclusion of a reflexive section is warranted.

Many teachers and researchers of translation have a background in industry, working as translation practitioners before or during their teaching careers. I likewise have a background in industry, but instead of translating, I worked for numerous years in both management, acting as a buyer of translation services in a major multinational enterprise, and then in sales at a large language services provider. The nature of the translation industry, with its tiered production networks (see Abdallah and Koskinen 2007, 877) therefore positioned me in a different part of the network than translator-practitioners and trained me to look at translation needs and processes from a different, higher level. This provides one explanation for why I decided to focus on non-translator users of MT rather than translators. It also explains my ease in initiating and carrying out collaboration with business partners

[19] MT story on patent professional: https://mt-stories.com/2019/09/27/machine-translation-for-patents/

as well as my decision in some studies to use interviewing, which was a process I had relied on in multiple contexts throughout my career.

My position in relation to the groups I studied varied. In Articles I and IV, which focused on users of an open internet tool, Multilizer's PDF Translator, I was an insider, as I had used similar tools on a regular basis. In fact, in Article IV I was also a participant in the research, since my own actions became part of the data. Conversely, I was decidedly an outsider in the study on patent professionals, having no experience with the genres, subject matters, or languages they dealt with daily. This no doubt affected how informants talked to me, perhaps forcing them to speak explicitly and explain things in more detail than they would need to with insiders. My outsider standing, however, was most distinct in Article V. The target users of the solutions discussed in the article were people who were barred from some of the information they needed in order to participate fully in their societies because they did not speak any of the dominant languages. It is important for speakers of well-supported languages, and especially for speakers of English such as myself, to recognize that we enjoy the privilege of almost always having access to the information we need. Throughout my research, I tried to remember the fact that I could not fully understand what life would be like without this privilege.

4.6 Ethical considerations during the dissertation

A project as extensive as this one necessarily involves a number of ethical issues to be planned for and resolved. In this section I describe some of the key issues encountered in my work.

4.6.1 Protection of personal data

During the research for this dissertaton, two laws governing the management of personal data went into effect: EU Regulation 2016/679[20], commonly known as the General Data Protection Regulation (GDPR) and Finland's Data Protection Act[21]. Personal data was collected for Articles I–IV. The data collection for Articles I and II was completed before GDPR and the Finnish Data Protection Act were in place. The data for Articles III and IV was predominantly collected while either the GDPR

[20] EU Regulation 2016/679 (GDPR): https://eur-lex.europa.eu/legal-content

[21] Finland's Data Protection Act: https://www.finlex.fi/en/laki/kaannokset/2018/en20181050.pdf

or both the GDPR and Finnish Data Protection Act were active. For this reason, and also due to my growing maturity as a researcher, the personal data protection for Articles III and IV was stronger than for Articles I and II. As an individual researcher, I acted as the data controller for all of the research in this dissertation.

My research did not involve inquiries about sensitive or very personal issues and to the best of my knowledge, none of the research participants belonged to a group whose participation would have rendered them vulnerable. Nevertheless, I did collect personal data for articles I-IV, and the methods I employed to protect personal data became increasingly robust as my research progressed. Participants in Article IV (which was the first chronologically) were informed that the data collected was used for research purposes only and that they would remain anonymous. The survey respondents in Article I were presented with information before starting the survey, including the aim of the research, information on the researchers, and again information that their answers would be used for research purposes only and that they would remain anonymous. For Articles II and III, I provided respondents with a more well-developed information sheet and consent form that was signed by all. In retrospect, only the participants in Article I remained anonymous, others were protected by pseudonymity.

4.6.2 Cooperation with companies

The cooperation with Multilizer Oy for Articles I and IV was initiated as a mutually beneficial exercise. The benefit for Multilizer would be a new approach to researching their users that would augment the information they regularly gathered through user surveys. The benefit for me would be access to their large base of users of an online MT tool for gisting. However, I would also face the challenge of ensuring I had independence when deciding on goals, methods and analyses.

Two aspects of the cooperation contributed to ensuring that the scope was not overly influenced by Multilizer. First, they had been studying their own users for years and one of their main motivations for cooperating with me was a need for new approaches to user studies. They wanted to see what innovations might be introduced through an academic endeavor. Therefore, they were motivated to leave me to define the scope and methods independently. An important second assurance I had on independence was the fact that they continued their own user studies throughout our collaboration. Any questions that they had that were specific to their own tool and business model, but were not interesting to academic research, could

therefore easily be covered in their own user surveys and I felt no pressure to include them. We jointly agreed that my research would not touch on market-sensitive data that the company did not otherwise share with the general public. Perhaps due to my background in business, I considered this to be a natural way of working.

The informants I interviewed for Articles II and III worked for private companies and governmental agencies. I asked informants from private companies whether their employers would have reservations with them speaking to me about their use of MT. As the informants assured me this was not an issue, I did not pursue requesting any further permission. As regards the governmental agency, I first spoke to a manager about identifying people for me to interview and that person found the informants for me. No formal agreements on cooperation were made beyond the consent forms signed by each participant.

One ethical issue that arose in my cooperation with business involved social media. I recruited participants for the study on patent professionals through my key informant, through participation in a patent professional conference, and through subsequent followups with people I met at that professional conference. After any discussion or exchange of mails with a new person, I requested to connect with them in LinkedIn. This was common practice for building networks in my past career, particularly in the five years I spent in sales prior to moving to academic work. I had successfully linked with five or six potential participants before I realized that, as I would be carrying out a qualitative study with a limited number of participants in a small field and a restricted geographical area, and as I did not have other people from that field as LinkedIn contacts, those links would provide an easy means for anyone to deduce the actual identity of participants. I corrected the situation by unlinking with those people and by sending an email to explain why I had done this. This issue had not been covered in any ethics course I had taken, but I made a point of introducing the topic in all subsequent ethics-related courses and events I participated in.

4.6.3 Researcher as participant

As previously mentioned, in Article IV I acted in a dual role, as both researcher and participant. Although I recognized this as a potential issue when planning the study, it was nevertheless a practical solution to conducting the interviews. Especially since one phase of the recruitment of interviewees involved contacting them online and starting interviews almost immediately, it would have been difficult to coordinate

with another person who would do the interviewing. The nature of the study also helped mitigate issues that might have arisen from my dual role. It was an inductive study in which I had not outlined initial hypotheses or specific behaviors I wanted to focus on. Therefore, I could not have tried to induce specific behavior from the participants. Despite this, a diary or better interview notes taken during the process could have lent robustness to the study and further reduced any negative effects from the dual role.

5 RESULTS, RESEARCH QUESTION 1: WHAT IS THE NATURE OF MT GISTING IN EACH OF THE CONTEXTS STUDIED?

I began this dissertation with a preliminary idea that certain elements in MT gisting contexts influence the use and reception of raw MT. To investigate this idea, I wanted to study real users of raw MT, but I needed to identify specific groups to focus on. As described in Chapter 3, I took advantage of all opportunities I found to locate different groups. I eventually explored three contexts in which MT gisting is currently taking place – the broad, very global use of one online MT tool; a closed ecosystem in which raw MT use is widespread and legitimate; and a communication situation conducted via MT-mediated communication – and one context in which MT gisting could be implemented, i.e., the context of underserved groups of people who might gain access to more of the information they need about the societies they live in through the use of MT.

At the beginning of the project, both my own knowledge of MT gisting, and the collective knowledge of the phenomenon that was available in scientific studies, were relatively small. Because of this, I needed an inductive approach and a broad research question to guide the work. Applying the question *What is the nature of MT gisting in this context?* to each context allowed me to explore aspects of MT gisting as I identified them and did not bind me to a predetermined idea of the types of factors I should look for. My findings on this research question are described in the following sections.

5.1 Online MT

The first of the dissertation articles that explored people's use of raw MT, Article I, focused on the users of an online MT tool, Multilizer's PDF Translator[22] (currently named Multilizer Document Translator). The research was conducted in collaboration with Multilizer and its CEO, Niko Papula, and I co-authored the

[22] PDF Translator (currently named Multilizer Document Translator): pdf.multilizer.com

article. In that study, our approach to answering the question *What is the nature of MT gisting in this context?* involved capturing a snapshot of the use of PDF Translator in a specific point of time. In the four-month period of November 1, 2017 to February 28, 2018, data was collected from usage logs as well as from a survey that was offered to every person who submitted a translation request during that period. The survey resulted in 1,579 responses. Two types of data were gathered in the study: general information on the users of PDF Translator and information related to the specific session of PDF Translator that triggered respondents to answer the survey. A discussion of each of these follows.

5.1.1 PDF Translator use and users

PDF Translator is a specific type of online MT tool. It translates entire documents instead of strings of text, and handles various formats, although at the time of the study, PDF was predominant. A free version offers a limited amount of translation (three pages at the time of the survey), after which users need to purchase translation in batches of pre-defined numbers of pages. The fact that PDF Translator is a paid service differentiates its users from the users of free online tools. Even if a person only uses the limited free service, we can assume they had a translation need that was not met by free online tools. Therefore, the users reflected in the study results were not a representative sample of all users of online MT tools, even though it was the broadest group explored in this dissertation.

The results of Article I revealed a geographically broad user base. Log information originated from 181 countries and territories, while the 1,579 survey responses represented 97 countries and territories. This is especially interesting considering that at the time, PDF Translator supported approximately 43 languages and the survey was offered in only 6 languages. Because PDF Translator has a very large user group based in Latin America and other Spanish-speaking countries, those regions were more prominent in both logs and survey responses than others.

The analysis of translation request logs from the study period revealed that English held a very strong position, with 85% of the translation requests during that time involving the source language of English. This was an expected result, as English has long been at the top of the list in similar studies. Also unsurprising was the predominance of the English–Spanish language pair, which comprised 47% of all requests. The popularity of this language pair has been reported in Yang and Lange (2003), Smith (2003), Gaspari and Hutchins (2007) and Turovsky (2016). The

top 10 source and target languages included many European languages, which was also expected. The largest surprise was Indonesian, which ranked fifth in target languages. A similar finding was reported in a blog by Google in 2016 (Turovsky), which reported that Indonesian was among the top five target languages requested in Google Translate.

When analyzing the demographic information from the survey, overall results were somewhat unsurprising. More males responded to the survey than females or people identifying as 'other' (68%, 32%, and 3%, respectively). The age group of 19 to 29-year-olds comprised 46% of all respondents, with older groups showing progressively fewer responses. The more interesting results were attained by comparing the demographics of responses from different language versions of the survey. This comparison revealed that male responses were more predominant in the French and Russian surveys (82–83%) and somewhat predominant in the Portuguese, Spanish and English (61–68%) surveys. But in the Indonesian survey, female responses outnumbered male 54% to 46%. Similar differences were revealed when comparing ages in different language surveys, with the Indonesian, Spanish and Portuguese responses showing a larger predominance of the younger age groups than the English, Russian and French responses. A final demographic finding was that the respondents to this survey were well educated: 50% reported having achieved at least a Bachelor's or vocational degree, with an additional 26% reporting some university or vocational studies.

5.1.2 Questions on what users translated in PDF Translator

A second purpose of the study was to inquire about the specific context of use of the MT gisting that was taking place when respondents answered the survey. Respondents were asked three questions concerning the document they had submitted for translation just before answering the survey. The first of the three questions asked why the respondent wanted to translate the document, aiming to identify if they meant to use the translation for assimilation or dissemination. This question was modelled after a similar question used by Gaspari (2007, 102–103) but included response choices that were more granular than Gaspari's. The results are depicted in Table 4. below.

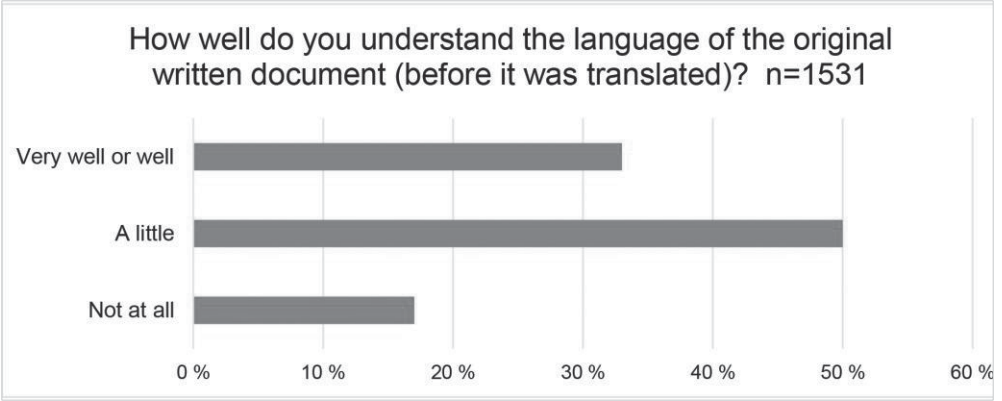
Table 4. Responses to the question: *Why did you translate the document [you submitted for translation]?* n=1528

Why did you translate the document?	% of responses
I wanted to understand it myself.	58%
I wanted to verify that I understood it myself.	18%
I wanted to translate it into my own language so that someone else can understand it.	14%
I wanted to translate it from my language into another language so that someone else can understand it.	6%
Some other reason (please specify)	4%

In both Gaspari (2007) and Article I of this dissertation, the first three responses were understood to indicate assimilation and the fourth, dissemination. However, in the categories of use proposed in this dissertation summary (Section 2.3.2), this would be seen differently. The fourth response would be considered to belong to the category of MT for gisting since there is no post-editing involved. Nevertheless, the granularity of the response choices provided interesting information on assimilation, for example, that people are translating information for others. Another interesting finding was that 18% wanted to verify that they had understood the document they translated. This indicates that, besides the translation, they had also read the source text.

A second question concerned how well respondents understood the source language of the document they translated. The results are shown in Figure 4. below.

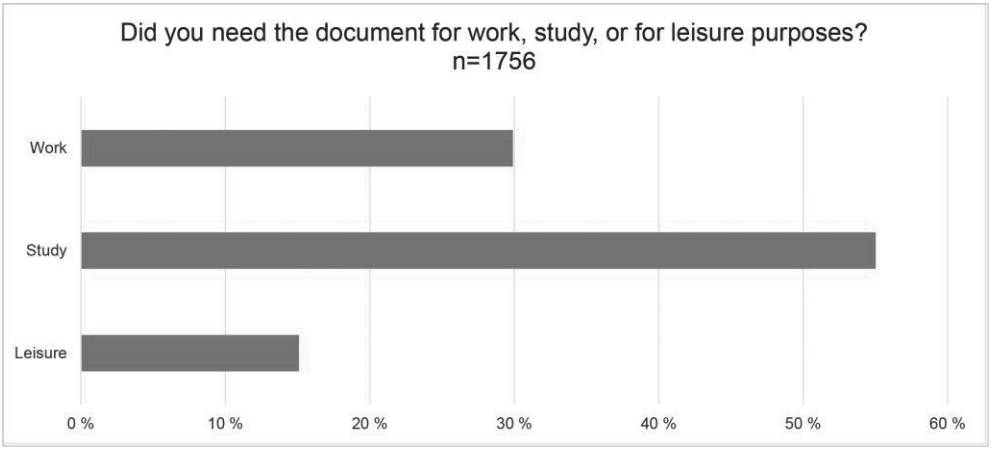
Figure 4. Responses to the question: *How well do you understand the language of the original written document (before it was translated)?*



The result that 83% of respondents claimed that they understood the source language a little, well, or very well was a clear indication that the users in this context were translating documents they already had some understanding of. Article I offered several possible explanations for this, including that respondents might have been testing the system, that they were cautious in trusting MT enough to use it with languages they did not have any proficiency in, or simply that people cannot locate documents they might want to translate if they do not know enough of the language to read the titles. Another consideration offered in the article was that perhaps this was an indication that people do not use raw MT in the same way they use human translation. Maybe people are using MT as a language tool that provides them with information that they can combine with other resources, such as their own basic knowledge of the source text, to arrive at an understanding of its meaning.

Article I's final question inquired about the area of life that the document submitted for translation concerned: study, work, or leisure. To the best of my knowledge, such a question had been asked only once in a previous study. Yamada et al. (2005) presented three surveys conducted in Japan in 2003–2005 which asked users if the purpose of their use of MT was personal or work. A general tendency toward using MT in the personal sphere was uncovered in those surveys. Article I expanded this question to include the choice of study. Respondents were also given the option of selecting more than one area in case a document had multiple purposes. The results of that question are depicted in Figure 5.

Figure 5. Responses to the question *Did you need the document for work, study, or for leisure purposes?*



Study was the most reported purpose for submitted documents, representing 63% of responses. Work was second with 33% and Leisure represented 18% of responses. Similarly to the demographics described in the study, this area also showed differences in the responses of the six different languages the survey was offered in. Further analyses of these differences can be found in Article I.

In summary, the most interesting findings on the broad user base of PDF Translator users were that people were using MT to both understand and verify their understanding of texts; that they were translating texts that were in languages they very often had some knowledge of; and finally, that MT was being used most often in the area of study, followed by work and leisure. Each of these would have provided a path for further research, but the next opportunity I found to study raw MT users involved the use of MT at work, namely, in the work of patent professionals.

5.2 Using MT at work: the case of patent professionals

The second group of people I studied for the dissertation were patent professionals working in the IPR field. These people use their expertise in patents to assist and guide others (internal or external groups) in IPR processes. Patent professionals frequently use MT gisting to review patent documents that are in languages they do not speak. Their main goal is to sift through large amounts of patent documents to identify the ones that are relevant to an IPR case they are involved in. Once they

have identified relevant documents, they may continue to use them in their raw machine-translated form, or they may have them translated by a human. More detailed descriptions of the environment and processes are available in Articles II and III.

The use of MT gisting in the IPR field is a diffuse and longstanding practice. Although the field has developed and employed MT in human translation processes for a much longer period, the widespread practice of using raw MT for gisting was estimated by Nurminen (2019, 32) to be approximately ten years old. Already in 2012, List suggested that "everyone involved with patent information" was using MT to access patent documents (List 2012, 193). When presented with the opportunity to study this group of people, I chose to employ qualitative, interpretive methods which involved interviewing nine Scandinavian patent professionals, most often at their workplaces. I thought that this would be an appropriate approach for a context that was entirely unfamiliar to me, and it would allow me to address the research question concerning the nature of MT gisting in the context. I eventually arrived at three conclusions. First, this environment constitutes an ecosystem and the use of MT in this ecosystem is transparent, governed by formal and informal guidelines, and enjoys a high level of legitimacy. Second, patent professionals employ a process of risk assessment and management in evaluating when and where to rely on raw MT. Finally, MT gisting in this ecosystem can be analyzed and understood through the concept of distributed cognition. Each of these is discussed further in the following sections.

5.2.1 MT gisting in an ecosystem

In the descriptions patent professionals gave of their MT gisting, it became evident that a key factor contributing to the feasibility of using raw MT was that the use occurred within an ecosystem that supported and accommodated it. As discussed in Section 4.3.2, an ecosystem involves a network of organizations involved in producing a product or service through both competition and cooperation, and the working environment described by the informants of my study fulfilled that description. Within that ecosystem, the use of raw MT was considered legitimate, tolerant of the risk inherent to MT gisting, and supported by technology. A further exploration of these themes follows.

5.2.1.1 Legitimacy of MT gisting

The first discovery I made in this study was that the use of MT in this ecosystem was legitimate. Before starting the research, I assumed that the use of MT would be an activity that neither the patent professionals themselves, nor the organizations they worked for, would discuss openly. In the first interviews, I asked informants repeatedly about their organizations' potential reactions to their descriptions of how they use MT. My inquiries were always met with surprised looks and assurances that it was fine and the whole industry was using MT in the same way. After some time, I finally understood the point I had been missing: in the IPR world, the practice of using raw MT was considered legitimate.

I arrived at this realization of legitimacy before I could attach it to a definition of the term. Merriam-Webster defines legitimacy as "conforming to recognized principles or accepted rules and standards"[23] while Cambridge defines it as: "the fact of being allowed by law or done according to the rules of an organization or activity."[24] Through careful analysis of the ecosystem and its use of raw MT, I derived three themes that described legitimacy for MT use in the context of this environment, which I included in Article II: "MT use was transparent, the boundaries of its legitimacy were documented and generally agreed upon by users, and its users had a relatively high level of 'MT literacy'" (Nurminen 2019, 39. For a further discussion on the concept of MT literacy, see Section 2.3.2). The first element of this description, transparency, was evidenced in interviews with informants as well as in the official guidelines from governmental patent offices, as described in the next paragraph. I derived the second element of legitimacy from the first one of transparency. According to the Cambridge definition above, legitimacy is constrained to the rules of a specific organization or activity. It would therefore be important that that legitimacy is recognized by participants as constrained to that context, and that there would be a shared idea of where, exactly, the boundaries of legitimacy lay. The third element of legitimacy was a collective MT literacy that was promoted and maintained by the members of the ecosystem.

After developing this initial idea on the components of legitimacy for the IPR context, I analyzed how these components were visible in my data from the first interviews with patent professionals, then added new questions for subsequent interviews. The transparency of the use of MT was consistently discussed by all informants of the study, including patent professionals working inside companies

[23] https://www.merriam-webster.com/dictionary/legitimate

[24] https://dictionary.cambridge.org/dictionary/english/legitimacy

that actively filed and prosecuted patents, patent professionals working as service providers to companies and inventors, and patent professionals working in governmental patent offices. This transparency was not limited to the people working most closely with the patent professionals. Rather, informants described how they shared machine-translated documents with their internal and external clients, and how transparency in the patenting process allowed even competitors to review, challenge, and offer new interpretations for machine-translated information (for a more detailed discussion, see Nurminen 2020, 108; 114–116).

To gather data on the second component of legitimacy, on where the boundaries of the practice lay, I asked all informants to describe situations in which it is not OK to use MT. There was general agreement that raw MT was not acceptable in legal settings; before discussing a case in court, relevant documents needed to be translated by a human. In all processes leading up to that, including discussions with governmental patenting offices which sometimes included outside stakeholders such as competitors, raw MT was permissible and the decision of using raw MT or human translation was left to patent professionals.

The third component of legitimacy, a collective MT literacy across the community, was developed and maintained in a variety of ways. First, formal guidelines on the use of raw MT are readily available. For example, the European Patent Office's guidelines for examination state that it is appropriate to rely on machine translation for documents in unfamiliar and non-official languages, and that "mere grammatical or syntactical errors, which have no impact on the possibility of understanding the content do not hinder its qualification as a translation" (EPO 2018, Part G, Chap IV-4). A second method for maintaining MT literacy is through training on the use of raw MT. Much of this occurs as part of the informal training of new employees, such as in this example given by PP5:

> But the actual reading of the texts is made by the researchers. And with a load of Chinese patents they have to read through machine translation, they have to be persuaded that they can actually see a meaning in those…typically I do just as with a small child, read aloud. And say how it is to be read and understood. And after such a procedure and some practice, most people realize that it's a useful tool.

The informants of the study demonstrated that a new competence should be added to the definition of MT literacy: the ability to assimilate information from raw machine-translated texts, which informants described learning and also teaching to others. For a more detailed discussion on this competence within the MT literacy framework, see Section 2.3.2.

5.2.1.2 Risk tolerance

Besides legitimacy, a second discovery I made about the IPR ecosystem that supported MT gisting was that the environment and the processes in it were tolerant of risk. To understand how MT gisting processes are tolerant of risk, it is first important to understand that work in the IPR context contains risks on many levels and in many processes. If a patent that infringes on an existing patent or patent application is erroneously granted, or a company builds the infrastructure to launch an innovation that is unknowingly protected by a patent, the consequences can be substantial. Such consequences might be financial, loss of time, or loss of goodwill within the marketplaces patent applicants operate in.

To help mitigate such consequences, the ecosystem has developed processes that are tolerant of risk. The patenting process itself is long and contains mechanisms that allow for detection and correction of oversights and errors. One such mechanism are third-party observation and opposition subprocesses, which allow for third parties to "bring to the attention of the authorities issues that did not emerge during the examination of the patent application" (PRH 2018, 19), therefore mitigating the risk of granting patents erroneously because a relevant patent document was overlooked by patent-granting authorities.

One of the most important actions patent professionals take to ensure that relevant documents are not overlooked is to ensure that all relevant patent documents are reviewed during the appropriate stages of IPR processes. A primary risk in their job is therefore missing a relevant document. This risk was brought up frequently in the study interviews, for example, in this description by informant PP3 (Nurminen 2019, 35)[25]:

> At work we talk about how most mistakes take place because someone overlooks a relevant patent…when a mistake happens, it is most likely to be caused by that […] putting it into the 'not interesting' pile is a risk.

Because they are responsible for all relevant patent documents regardless of language, one of the most important tools patent professionals use in accessing and identifying relevant documents is raw MT. And herein lies the second important point in understanding risk in this ecosystem. Raw MT is used to mitigate the higher-level risk of missing relevant information. A definition of risk mitigation by Pym and Matsushita (2018, 2) states that mitigation is the attempt to protect against one risk

[25] When an informant's quote was already published in one of the dissertation's articles, a reference is given. When the quote is published for the first time in this summary, no reference is given.

by incurring a second risk, which does not remove the first risk but minimizes its consequences. Accordingly, patent professionals take the risk of relying on raw MT to mitigate the larger risk of neglecting to consider relevant patent documents.

Despite being a mitigation, the use of raw MT introduces other risks. For example, even if relevant patent documents are identified, they can be misunderstood due to poor translations. An awareness of this risk was discernible in the process that patent professionals described using to decide on whether to rely on raw MT or order human translations of documents. A discussion of this process follows in Section 5.2.2.

However, the tolerance for risk built into higher-level IPR processes also provides tolerance for the lower-level process of using raw MT gisting to identify and read relevant patent documents. For example, the long patenting processes described above also provides a time and place for identifying and correcting misunderstandings in machine-translated texts, as described by PP6 (Nurminen 2019, 38):

> Well of course you can get the wrong impression of the subject matter in the document, but I don't see that it's a really big risk because the patent application process is a long process, so if my interpretation of some kind of document based on the machine translation is wrong, I can change my mind later, if I see it. It takes usually over two years to get a patent so we get the answer from the applicant and we probably write another office action and then the applicant replies again, so it's a conversation. So during the process there's many times when these things can be dealt with.

The risk tolerance that MT gisting is afforded because of being embedded in higher-level processes that are risk tolerant illustrates a paradox in MT gisting. Much discussion has been held about the risks involved in relying on raw MT in risky and critical contexts. However, in the IPR example, it is exactly this inherent risk in the environment that has prompted organizations to build tolerance for risk into their systems and processes. This implies that organizations that work in high-risk circumstances may, in some cases, be better equipped to tolerate the risk involved with using raw MT.

A final way in which risk tolerance was visible in the IPR ecosystem was the familiarity patent professionals displayed toward working within the constraints of risk. The topic of risk was raised often as informants described the environment of patenting work and the risk involved, not only in MT gisting but also in their other processes. In the following example, PP2 illustrated the various risk points in

searching for and reviewing patent documents, with the first being the search for patents and the second being reading and trusting raw MT:

> And if there's something crucial, high risky thing that you throw into the trash. Then it happens. And then you're…at least you get some kind of confidence, you get some kind of result, and then you just…the business makes the decision based on all other things and patents as well, but anyway, you do your best and something is maybe always left out. Sometimes you get away with it, sometimes maybe you don't, but at least you have done your homework and then if something happens then you just take care of it. But yes, there is risk when you make decisions based on something, some patents you leave out, some things you take in, and you do it based on machine translations and…that's life.

Article II suggested that the informants' high tolerance of risk might be the result of their considerable experience in the field, their organizations' willingness to accept the inherent risks of IPR work, and an acknowledgement that reliance on raw MT was simply necessary, since it would be impossible to fulfill all translation needs through human translation.

5.2.2 Using risk assessment in decision making

The second conclusion of this study was that patent professionals' decisions to rely on raw MT relied on assessments of risk that reflected the principles of risk management. At the beginning of the study, one of the various aspects I was interested in exploring was trust. I wanted to know how patent professionals viewed their trust in the practice of using MT, including when it is wise to trust MT and when it is not, and how they assessed their trust of specific machine-translated texts. In my interviews I asked a variety of questions surrounding the concept of trust, including questions on understanding, legitimacy, decision making, and risk.

During the data analysis process, I came to realize that my original assumption on the primacy of the concept of trust was incorrect. I had assumed that the main issue in MT gisting for patent professionals would be trust, and one component of trust was risk. What I came to realize was that the *main* issue for patent professionals was not trust but risk, and that trust was a component of that risk. The consideration of risk played an important role in much of their work, and in making the important decision of whether to rely on raw MT or opt for human translating, these professionals displayed a conscious decision-making process that weighed risks, consequences and benefits.

In this decision, the arguments in favor of relying on raw MT were first, a substantial savings of time and costs, since patent professionals could access raw machine translations of documents in a manner of seconds. The second argument for using raw MT was the patent professionals' trust that they had a sufficient understanding of the raw MT to be able to use it. This understanding was not the simple result of the patent professional's own interpretation of the raw MT. Instead, the IPR context contained numerous factors that contributed to a process of building an understanding. Examples of such factors were the patent professionals' familiarity with the patent genre, the familiarity with the subject matter that technical experts had, their knowledge of the source language, their access to multimodal elements in the source document, alternative MT outputs, reliance on other patent professionals, and the affordances listed in the previous section: legitimacy and risk tolerance.

On the other side of the equation were several arguments in favor of ordering a human translation of a document. First, patent professionals considered the assumed relevance of a document to the IPR process in which the machine-translated document would be used. If the document appeared to be highly relevant, it might be sent for human translation immediately. The second point of consideration was the riskiness of the IPR process in which the MT would be used and the potential consequences that might result from a misunderstanding of the text. This was voiced by informant PP4 (Nurminen 2019, 36):

> if the context is clear then it's OK as I see it, I trust the machine translations enough, but sometimes when we are in borderline decisions it's required to have a proper human translation…So it's more a question of the uncertainty margin of the translation with respect to what we are deciding.

Informant PP8 summarized the sentiment concisely: "The more important the decision, the less you do the decision based only on machine translation" (ibid.).

5.2.3 MT gisting and distributed cognition

As described in Section 4.3.2, Articles II and III relied on the same data set, which was analyzed using two separate methods. I had already completed the first version of Article II when I began to consider whether situated cognition might be a useful lens through which MT gisting could be examined. I already had a data set that presented a deep description of how an ecosystem used raw MT and, based on that data, I had concluded that one possible way to understand the phenomenon would

be through the concept of risk management. For Article III, I analyzed the data again through the framework of distributed cognition to see if that would be an alternative way to view the phenomenon. For a discussion on definitions and theoretical work on distributed cognition, refer to Section 3.6.

In the analysis of the distributed cognition at play in the IPR ecosystem, I first borrowed the idea of the extended mind as outlined by Clark and Chalmers (1998, 8–9). According to this concept, the patent professional and the MT tool form a *coupled system* comprised of a human component and an external system. The participation of the external system, the MT tool, is "active" in that its removal would greatly hinder the performance of the coupled system (ibid., 9). In other words, if the MT engine were removed, the patent professional's ability to comprehend the original text would decrease dramatically.

Beyond this coupled system, the analysis identified four elements in the IPR context, including both artifacts and other people, that form a network that patent professionals interact with to achieve a necessary understanding of raw MT. First, in addition to the machine-translated version of the patent document they need to read, they also accessed the original document in its original language. Depending on the source language and their own language competences, they relied entirely on the original texts or they read both the original and the raw MT and combined their grasp of both to achieve an understanding of the text. Additionally, patent professionals often accessed the non-textual components of the original documents, such as drawings and chemical and mathematical formulas, to augment their understanding.

Another important element in their network that helped patent professionals achieve an understanding of raw MT were the technical experts who were behind the inventions that required patenting. Whereas the patent professionals had the best expertise on the patent genre, it was the technical experts who knew the subject matter best. The study's informants described how it was the combination of these competences that often led to enhanced understanding of raw MT, and this was gained through formal and informal discussions. This was described by PP1 as follows (Nurminen 2020, 112):

> But many times we are in a meeting. I have the people there. I have [the tool] open and we are reading from it [...] and then when there's a German or French or Chinese text, we realized that oh, that's what we're dealing with, and the button is there and ready, Translate, and then it does it.

The third element in the network that patent professionals interacted with was an alternative MT tool. If their primary MT tool did not produce satisfactory output,

they put the same document into another MT tool. Often these were MT tools provided by individual governmental patenting agencies, but informants mentioned also using Google Translate on occasion.

The final element in patent professionals' network that they interacted with to augment their understanding of raw MT was a large network of other stakeholders. Patenting and other IPR processes often involve large, and usually international, networks of IPR service providers, governmental patenting agencies, and private companies who all have a stake in the outcome of the process. Patent professionals can also interact with this network of people when they need help with understanding a patent document. For example, if only one detail of a machine-translated patent document is not understood, they might be able to access someone who understands the original document for an explanation of that one detail.

Article III concluded that distributed cognition was a useful and appropriate way to analyze and understand the MT gisting used by patent professionals. The framework was instrumental in highlighting the interactive nature of MT gisting in this context. A second conclusion was that the application of distributed cognition led to a more nuanced understanding of the use and reception of raw MT in the IPR ecosystem.

5.3 MT-mediated interviewing in research

Chronologically, the first investigation I made into MT gisting was a study conducted with four users of Multilizer's PDF Translator tool, the same tool whose users I surveyed for Article I. In this study, the research question *What is the nature of MT gisting in this context?* focused not on users' reports of their own MT use, but on MT-mediated communication itself. The purpose of the study was to pilot the use of MT-mediated communication for data gathering in research. The expected results were that it would prove to be promising enough to warrant further study, which it did, and that the study would uncover contextual factors that were found to affect the MT-mediated communication of the interviews. In fact, this was the only study of the dissertation with a main goal that was tied to RQ2, concerning contextual factors that affect the use and reception of raw MT.

The participants in the interviews were four Spanish-speaking users of PDF Translator and myself, as I acted as the interviewer. The interviews were conducted

using Skype Translator, which at the time was a "preview" version[26]. This tool had an instant messaging or chat function into which MT was integrated. I specifically recruited Spanish-speaking participants because the English–Spanish language pair had long been recognized as one that yields high-quality MT results and also because I had a basic proficiency in Spanish and guessed that it might prove useful. The data that was analyzed included the communications throughout the process, my own observations as a participant, and transcripts from the interviews. As mentioned previously, it was the communication that took place in the interviews that was analyzed, not the content of the participants' respondes to interview questions.

The study uncovered seven "considerations" for using MT-mediated communication in data-gathering interviews (Nurminen 2016, 74). Two of the seven considerations concerned aspects of interviewing via chat that were not specific to the use of MT in interviewing. First, when interviewing geographically displersed informants, time zones need to be considered. Second, while participating in the online chat-based interviews, informants appeared to be multitasking, although this did not lead to overly long lags in responding.

The remaining five considerations did relate specifically to the use of MT-mediated communication. The first of these concerned the time required to conduct MT-mediated interviews, which the study concluded to be relatively high in comparison to the time needed for face-to-face or monolingual interviews. A second was the user experience reported by the informants. When asked about the experience of using MT-mediated communication at the end of each interview, all of the informants reported a positive experience, responding that they felt that the method worked and that they understood and were understood.

Three final considerations related to the contextual factors that influence the use and reception of raw MT that were also uncovered in other parts of the dissertation as input to RQ2. They are therefore covered in Chapter 6 as follows: technology (Section 6.3.1), understanding and negotiation for meaning (Section 6.2.4), and competence in the language of the communication partner and tendency to adapt messages for better MT output (Sections 6.1.1 and 6.2.3, respectively). A final aspect that was mentioned briefly but not identified as a consideration in the article, involved familiarity with subject matter. This is discussed further in Section 6.1.2

[26] Currently Skype Translator's text translator: https://www.skype.com/en/features/skype-translator/

5.4 MT as a means for increasing accessibility to information

The final area I explored in the dissertation involved how MT could be used to provide better access to information for groups of people who do not currently enjoy such access. These people can be blocked from important information due to a lack of competence in the languages in which such information is published, and this can hinder their full participation in the societies they live in. Affected groups include people who have recently moved to a new country, refugees and asylum seekers, and people in crisis situations. In this area, the research question *What is the nature of MT gisting in this context?* did not directly concern actual raw MT users. Rather, it involved an analysis of research and projects focused on defining contexts in which MT could be used.

Article V was different from the others in this dissertation in that it addressed the use of MT for post-editing alongside its use for gisting. This was necessary in the contexts being studied, since all types of translation and levels of human involvement needed to be evaluated against each other for different information types. It also brought important viewpoints to the dissertation as a whole. It was a context that concerned a different type of MT user than in the other studies, which often involved people who had agency in their use of MT. Either it was a specific choice they made or they were provided training and support in its use. When MT is used for increasing accessibility to information, it reaches users who do not always enjoy such agency and have little choice in how they access information. Perhaps the most important new viewpoint this context brought to the dissertation, however, was an ethical one. First, it introduced the idea that employing MT for humanitarian purposes should be part of the discussion on ethics and MT. Second, it raised questions that need to be considered when deciding between using raw and post-edited MT, therefore highlighting important ethical considerations in MT gisting.

In Article V, a number of projects that have planned or implemented MT to improve accessibility were examined. The first group focused on improving civic participation, with example projects devoted to public service MT (Vasiljevs et al. 2014; Jönsson 2016; Miyata et al. 2015). The second group aimed to improve access to health and safety information in the areas of public health (see, for example, Aymerich and Camelo 2009; Kirchhoff et al. 2011; Birch et al. 2018) and public safety, including safety in crisis situations (Federici and O'Brien 2020). The third group focused on access to culture and media and included projects focused on subtitling (see Matamala et al. 2015; Melero, Oliver and Badia 2006). Most of the efforts in projects focused on MT + post-editing, although MT gisting was included

in the scope of some. A common theme was that projects were managed as consortia between stakeholders that included governments, academic institutions, and private companies. This is a promising development, since earlier research predicted that MT solutions focusing on humanitarian purposes might only interest governments (see Carbonell et al. 2006; Somers 1997). Another theme was that studies on projects that were being planned and piloted greatly outnumbered studies on up-and-running systems and their users.

This particular use case for raw MT is a promising one. Increased mobility and movement in the world have resulted in a growing understanding of the problems associated with language barriers and a desire to find solutions for those problems. In research, the idea of using MT to lower these barriers was cited as one of the early motivations for the development of the technology (Hutchins 1986, 15). More recently it has been included in discussions on ethics and MT (see Kenny, Moorkens and do Carmo 2020), which is part of a larger research focus on ethics in artificial intelligence. Using MT to increase accessibility fulfills one component of ethics as envisioned in this arena, namely, a prioritization of the humanitarian benefits of technology (IEEE 2017, 23). Additionally, this application of MT is highly effective in highlighting the innovation in human communication that MT offers.

Some promising signs for future projects can be detected in this area. The first is a recent research focus on developing solutions for the languages that are needed by many groups in accessibility scenarios. These languages tend to suffer from a lack of the digital resources needed to develop MT solutions (they are therefore called 'low-resource' languages), and a lack of attention because they do not have broad commercial interest or support. A second promising sign is an increase in the presence of the targeted user in these projects. Recent projects funded by the EU, for example in the Horizon 2020 program, have included a component on the targeted users of systems to be developed[27].

However, there are also challenges in applying MT to improve accessibility. The use of raw MT for certain types of information, for example health and medical information, can be considered too risky. As a result, much of the benefit is focused on MT and post-editing. It is more difficult to achieve very large increases in the amount of information delivered in other languages through post-editing than it would be to deliver information in its raw MT form, indicating that some of the potential of MT is not being realized. Also, as discussed in Section 5.2.1, instead of

[27] Examples of such projects include GoURMET: https://gourmet-project.eu/ and MeMAD: https://memad.eu/

removing risk, it might also be possible to make processes more tolerant of risk. However, this aspect was not considered in the literature analyzed for the study.

A second challenge in improving accessibility through MT, also investigated in Article V, is that there are ethical issues to be addressed. First, much of the information that would be included in accessibility efforts is safety-critical and therefore, not considered suitable for raw MT (Kirchhoff et al. 2011; Das et al. 2019; Turner et al. 2019). Second, the view of the targeted users of machine-translated information must be considered. Past research indicates that users' views on the acceptability of receiving information as raw MT can be influenced by the purpose they have for the information (see Bowker 2009; Bowker and Buitrago Ciro 2015; Birch et al. 2018). A suggested way of ensuring targeted users' views are considered is to involve them directly in research projects (see Bowker and Buitrago Ciro 2015; O'Mathúna et al. 2019, 8).

A final challenge is that projects in this area tend to be conducted through collaboration between researchers, governmental bodies and private industries. This might prove more challenging to maintain beyond development and into implementation phases. Indeed, one observation from Article V was that finding information on projects that are being planned or implemented was easier than finding information on established systems and their users. It was unclear whether this was caused by disinterest in the eventual users of systems or by projects being discontinued after initial system development.

6 RESULTS, RESEARCH QUESTION 2: CONTEXTUAL FACTORS THAT INFLUENCED THE USE AND RECEPTION OF RAW MT

As stated in the introduction, I started this dissertation with the idea that there are specific factors in the context of use of MT gisting that influence how people use and receive raw MT. This evolved into Research Question 2: What factors in the contexts of people's use of raw MT influence their use and reception of it? A version of this question was among the research questions for Article IV, which aimed to uncover factors that should be considered when using MT-mediated communication. However, it was not the main research question of any of the remaining four articles. Despite this, it was a common thread that ran throughout the project and each article contributed to RQ2.

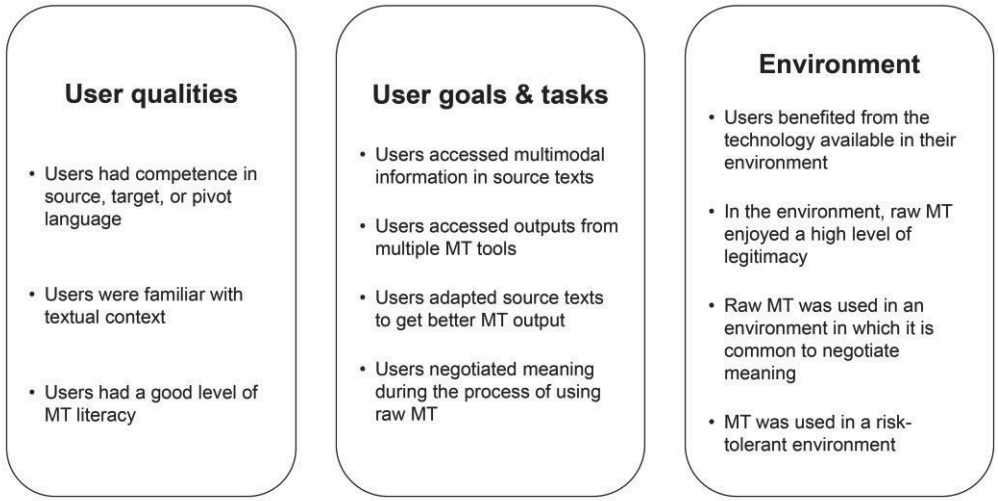
Eventually my research uncovered eleven contextual factors that were reported or observed to influence participants' reception of raw MT in some way. I refer to these factors as *contextual influences* in this dissertation summary. As most of the research for this dissertation was inductive and qualitative, its scope did not include empirical testing of the effects of these factors on reception. Therefore, the results of the research are restricted to reports and observations. However, the majority of the factors have also been identified in past studies, and some of those involved empirical testing. Moreover, all of the factors listed should be explored further and studied empirically in future research to attain a more precise understanding of their influence on raw MT reception.

The contextual influences discussed here were revealed individually during the dissertation work. I began to systematically document them early in the process but did not follow any existing typology or model for searching for, identifying, or documenting them. As the list of factors I was documenting grew, however, it became necessary to categorize them in order to more easily conceptualize and discuss them. I decided to return to the ISO 9241-210:2019[28] standard's definition

[28] The ISO standard I used at the beginning of this thesis was ISO 9241-210: 2010, not 2019. However, I switched to the 2019 version when it became available. To maintain consistent referencing throughout this dissertation summary, I updated all references to reflect the 2019 version of the standard.

of *context of use* (ISO 2019, 12) and its categories of user qualities, goals and tasks of the users, and environment. As described in Section 2.2, I had relied on this definition early in the PhD project to help me outline the areas of context I wanted to study and to then identify specific questions to be used in surveys and interviews. The categorization offered in the definition proved to work well for structuring the list of factors and I adopted it. It is possible that, as research in the area grows and new factors are found, this categorization may change or be discarded. But for the purposes of this dissertation and starting a broader discussion on the role of context in raw MT reception, this categorization is appropriate. It is displayed below.

Figure 6. Contextual factors that influenced the use and reception of raw MT, categorized according to the definition of context of use given in ISO standard 9241-210:2019



In the following sections, I further explore each of the contextual influences that was identified and reported on in at least one of the studies of this dissertation. Since most of these factors have been examined in past research as well, I have also included a discussion about related work on the same contextual factor. Within each category, studies are presented thematically and in chronological order from the oldest to the most recent.

This chapter does not aim to present an exhaustive list of contextual influences. It rather intends to provide a starting point for future exploration of the phenomenon and for empirical testing. As we learn more about how people use MT

gisting, as the ways people use raw MT develop, and as new groups of people start to use raw MT, this list can and should expand. The factors uncovered in this dissertation are introduced in Table 5. and discussed individually in Sections 6.1.1 through 6.3.3.

Table 5. Contextual influences in raw MT reception

Section	Factor	Description	Articles
User qualities			
6.1.1	Users had competence in source, target, or pivot language	The knowledge (anything from a little to extensive) users had of the source language, the target language, or a pivot language involved in the translation positively influenced their use and reception of raw MT.	I, II, III, IV
6.1.2	Users were familiar with textual context	Users' background knowledge of the source text (subject matter and genre) or the topic of conversation positively influenced their use and reception of raw MT.	I, II, III
6.1.3	Users had a good level of MT literacy	Users' understanding of how MT works and what types of limitations it has influenced their use and reception of raw MT.	II
Goals and tasks			
6.2.1	Users accessed multimodal information in source texts them	Users augmented their understanding of the raw MT by also accessing multimodal information available in the source text, for example, images, chemical formulas, or mathematical equations	II, III
6.2.2	Users accessed outputs from different MT tools	Users put the same text into different MT tools and used the outputs from both, or they were automatically offered two different outputs. They either compared the translations and used the one they thought was better, or they combined their understanding of each to arrive at an overall understanding.	II, III
6.2.3	Users adapted source texts to get better MT output	Users adapted their own messages to achieve better MT output, sometimes giving different inputs until satisfied with the output. Information producers adapted source texts (for example, through Controlled Language) to get better MT output.	IV
6.2.4	Users negotiated meaning during the process of using raw MT	In the process of using MT, meaning was negotiated. Participants in MT-mediated communication verified their understanding and asked questions to maintain common ground. Although past research has not identified this tendency in conjunction with MT and static texts, this dissertation uncovered the practice in the context of patent professionals' work.	II, III, IV

Environment			
6.3.1	Users benefited from the technology available in their environment	Users did not rely solely on MT tools, but also on the higher-level software or platforms that the MT tools might be embedded in, additional features offered by the MT tools, and other tools available to them.	II, III, IV
6.3.2	In the environment in which it was used, raw MT enjoyed a high level of legitimacy	MT use was transparent: stakeholders were aware that it was being used and machine-translated information was clearly marked as such. The boundaries of the legitimacy of MT were generally agreed upon by stakeholders. There was a fairly high level of collective MT literacy among the MT users in the context.	II, V
6.3.3	Raw MT was used in an environment in which negotiation for meaning is common	Raw MT was used in an environment in which meaning tended to be negotiated. Therefore, the use of MT and the meaning negotiation associated with it was better tolerated.	II, III
6.3.4	Raw MT was used in risk-tolerant environment	An environment that tolerates risk well can provide tolerance for the risk involved in using raw MT. It can also result in raw MT users who are accustomed to working within the constraints of risk. Conversely, in contexts that are not tolerant of risk, for example contexts in which content related to health or medicine are relied on, the use of raw MT is problematic.	II, V

6.1 User qualities

User qualities are the characteristics that a person brings into the situation in which they are using raw MT that can be considered to influence their behavior and their reception of raw MT content. According to ISO 9241-210, characteristics that influence the use of a system include "knowledge, skill, experience, education, training, physical attributes, habits, preferences and capabilities" (ISO 2019, 31). Henisz-Dostert (1979, 202) describes what this means in the case of MT gisting: "Once again, we are confronted with the contribution that the user himself [sic] brings to his interaction with machine translation, whether it be his familiarity with the subject matter, his use of context, his correcting the imperfections mentally, or his accommodation to the peculiarities of the texts."

The individual factors concerning user qualities that were uncovered in this dissertation are listed in Table 6. on the next page.

Table 6. User qualities that influenced raw MT use and reception

Factor	Section in which factor discussed
Users had competence in source, target, or pivot language	Section 6.1.1
Users were familiar with textual context	Section 6.1.2
Users had a good level of MT literacy	Section 6.1.3

6.1.1 Users had competence in source, target, or pivot language

The influence of a raw MT user's proficiency in the languages involved in MT gisting, including competence in the source, target, or a pivot language, was identified and discussed in four of the five articles of this dissertation (Articles I–IV) as well as in past research. Studies revealed evidence that proficiency, even low proficiency, in the source or a pivot language can have a positive influence on MT gisting. This finding offers new insight to the common assumption that the prevalent user group for MT would be people with no proficiency in other languages. Contrary to that assumption, the evidence discussed in this section suggests that MT can augment already existing language competences and would therefore be useful for people who have some level of proficiency in the languages they want to translate texts from. This does not, however, mean that MT would not also be useful to people with no competences in the source languages. For example, although the patent professionals in Articles II and III described benefitting from source language proficiency when reading patent documents translated from languages that they had competence in, they nevertheless also commonly read documents that were translated from languages they had no competence in. A few studies mention the important point that the benefits from source language competence are best reached when MT tools make it easy for users to see both source and target texts.

The first study of this dissertation to identify language competence as being influential in MT-mediated communication was Article IV, which was chronologically the first of this disseration. The article involved interviews with four informants that were conducted using a chat application with integrated MT. The interviewer spoke English but had a basic level of proficiency in Spanish. The informants were all Spanish speakers, and three of the four reported having a basic level of proficiency in English. The influence of this competence on the communication in the interviews was evidenced primarily in a tendency for participants to adapt their own messages to get better MT output for their

communication partner, which is discussed more thoroughly in Section 6.2.3. A second demonstration of the influence of language competence was in one informant's comment on their experience with using MT-mediated communication: "Very good, because it allows me to review the complete translation in the original language and in my language" (Nurminen 2016, 80). This introduced an important consideration concerning how language competence can positively influence MT gisting; namely, that the benefits of language competence are maximized when people are shown both source and the target texts (ibid., 81).

The influence of language in that first study led me to include a question on source language competence in my next study, which was a broad survey of users of Multilizer's PDF Translator tool. This survey led to Article I in the dissertation. In one section of the survey, respondents were asked about the document they had submitted to the tool immediately before being invited to take the survey. One of the questions asked was, "How well did you understand the language of the original written document (before it was translated)?" (Nurminen and Papula 2018, 206). The results of the survey revealed that a high percentage of respondents, 84%, reported they understood the source text of the document they translated very well, well, or a little. Only a small percentage had translated a document from a language they did not know at all. This indicated that the respondents tended to employ raw MT with texts that were in languages they already had some proficiency in. For more on this, see Section 5.1.2.

In Articles II and III, patent professionals reported that their competences in various languages helped them understand patent documents in those languages. Informants described reading the source text in its original language and the machine translation of that text side-by-side, then arriving at meaning by combining the two. The process was described by PP4 as follows (Nurminen 2019, 38):

> The complementarity of understanding the structure of the language better than the machine, and the machine understanding more words than I do, is a good complementarity.

Conversely, informants commented on the effects of reading documents from the most-translated languages of Chinese, Japanese and Korean, which most had no competence in. Due to the large number of patents being produced by the respective countries, patent professionals need to read such translations frequently. All informants in the study raised this issue independently, commenting on the negative influence of reading content from those languages, with problems caused by both a

lack of quality in the translations and not having competence in the languages in question.

6.1.1.1 Related work

Several past researchers have hypothesized that competence in the source language might positively influence MT gisting. Hutchins and Somers (1992, 157) predicted that, "[s]uch raw MT output may be even more useful to someone knowing the grammatical rudiments of the source language but not enough vocabulary to read texts fluently." A usability evaluation on several MT systems carried out by Gaspari (2004) included a component studying which systems made it easy for users to do "parallel browsing" of a web page and its machine translation (ibid., 76). This component stemmed from the author's assumption that such a feature could be helpful for users who had some, even limited, competence in the source language. Conversely, people with high proficiency in both source and target languages were hypothesized by Gaspari (2007) and Gaspari and Hutchins (2007) to benefit from translating individual words and short phrases with online MT, as "their bilingual knowledge would enable them to be in control of the process and vet the acceptability and correctness of the output" (Gaspari 2007, 116). For people with less familiarity with the languages involved in their online MT use, Gaspari and Hutchins (2007) concluded that electronic dictionaries would be a better choice.

In addition to these hypotheses on the effects of language competence on MT gisting, a number of empirical papers have investigated how competence in source, target, and pivot languages influences the reception of raw MT. Henisz-Dostert's (1979) survey of scientists who adopted raw MT for reading scientific articles found that the scientists' lack of competence in the source language of Russian might have contributed to a positive attitude towards MT, since the scientists truly needed translations to access the texts. The study also concluded that no source language knowledge was required to be able to use the machine-translated content (Henisz-Dostert 1979, 173). Informants in Birch et al. (2018) reported that competence in the source language of English would be a way to verify understanding of MT. People who understood a little English could use the source text together with the machine-translated text "to ensure their understanding was correct by comparing the two versions" (ibid., 26).

Four articles have investigated whether MT was more beneficial to people with lower competence in the source language than those with a higher level of competence. Both Fuji et al. (2001) and Morland (2002) found some evidence that

MT users with lower levels of competence in the source language of English benefitted more from MT than MT users with a higher level of proficiency in English. Fuji and colleagues emphasized that the benefit to those with lower proficiency in English occurred only when those people were shown both the source text and the MT. Similarly, Calefato et al. (2012) studied the impact of source language knowledge on participants in multilingual chat meetings conducted via MT-mediated communication. Two groups, one with higher proficiency and the other with lower proficiency in English, were asked about their satisfaction and comfort during the multilingual chats. Contrary to the first two studies mentioned above, Calefato and colleagues did not discern a difference in the benefits of MT as perceived by the different groups. Finally, a study on the use of MT by nurses in Japan found that nurses with a moderate proficiency in English benefited more from MT than those with limited proficiency (Anazawa, Ishikawa and Takahiro 2013, 26). Taken together, these results are inconclusive on the question of how varying levels of proficiency in the source language affect raw MT reception.

In addition to examining the effect of source language competence on scientists' use of raw MT, Henisz-Dostert (1979) also investigated the effect of target language proficiency. While only 34% of the informants in the study spoke English as their native language, all were reading scientific articles machine-translated from Russian to English. The study aimed to determine whether proficiency in the target language led to a higher tolerance for imperfections in the MT output. Although no definite conclusions were reached, "[t]he general impression was that the greater the familiarity, the more tolerance there was" (ibid., 172).

A further aspect of language competence that has been studied in relation to MT gisting involves competence in a pivot language. In a study by Ogura et al. (2004), multinational informants performed a joint task, communicating through MT in a bulletin board-type communication tool between their languages of Chinese, Korean and Japanese. The system had a feature that showed participants the initial translations from their own languages into the pivot language of English. It then allowed them to adapt their own messages until they were satisfied with the English results, after which they approved the messages which were then translated into the subsequent languages. The study found that participants' competence in the pivot language of English contributed to improved communication among the group because, "…all users, regardless of their nationality, often used the English translation to better understand the original messages" (Ogura et al. 2004, 600).

As was noted in Article IV, in order for users of raw MT to attain the most benefit from their source-language competence, they should have easy access to both source

and target texts. The importance of easy access to the source text has been noted in past research as well, for example in Morland (2002), Gaspari (2004), and Tinsley et al. (2012). Shigenobu (2007) also recognized the benefit of showing users back-translations of their own messages next to the original messages. This helped them be aware of how their messages were being translated and facilitated their adaptation of those messages to get better MT output (for more on adaptation, see Section 6.2.3).

6.1.2 Users were familiar with textual context

A second factor that was observed or reported to influence people's use of MT is their background knowledge of, or familiarity with, the context of a text. In this dissertation the *textual context* is defined as the subject matter and genre of a text. Therefore, textual context is one part of the entire context of use that this dissertation is concerned with. Familiarity with genre and subject matter were identified as having an influence on raw MT reception in Articles II, III and IV as well as in various previous studies. Collectively, the studies presented in this section displayed evidence that familiarity with subject matter and genre contributes positively to the reception of raw MT.

The positive influence of familiarity with subject matter first appeared in Article IV, which was chronologically the first of the dissertation, but only in a brief comment on a point of misunderstanding: "The first instance of lack of understanding involved an acronym. Fortunately, I happen to know it, but I asked for confirmation to make sure" (Nurminen 2016, 78). It was my professional background and familiarity with the informant's occupation that allowed me to guess the meaning of the acronym and continue that line of conversation. The fact that I gave the incident such a small mention in Article IV clearly shows that, at the time the article was written, I had not yet realized the importance of that part of the exchange and had yet to analyze the influence of contextual familiarity on reception.

In Articles II and III, patent professionals gave detailed accounts of how familiarity with the two textual context aspects of genre and subject matter positively affected the success with which they worked with raw machine translations of patent documentation. Informants reported that their own competence in the patent genre helped them to understand the documents: "there's a certain structure and there's a certain format that they're in. Then it's in a way easier to follow" (PP2 in Nurminen 2020, 112). However, they also often pointed out that the inventors behind the

patents they were working with were the real technical experts who understood the subject matter thoroughly. They indicated that it was the combination of their own expertise in the genre and the technical experts' knowledge of the subject matter that often proved to be a key factor in understanding raw MT. This is illustrated in the following comment from informant PP1 (Nurminen 2020, 113):

> So the system goes that way that, when we got the search results I send them out to the technical experts. They read them first by themselves. And they pick out those that they are worried about or where they want to have more information, where they're not sure what it really means. So then we have a meeting and we go through them together. So I can tell them what it really means and how to read it.

6.1.2.1 Related work

The relationship between textual context and readers of various types of texts has been discussed for years[29], so it should not be surprising that it would also be a point of interest when people are reading machine-translated information. As early as 1992, researchers suggested that familiarity with subject matter might be helpful when reading machine-translated texts. For example, Hutchins and Somers (1992, 157) proposed that:

> Experts in scientific fields need access to current documentation in languages they cannot read, e.g. reports on space technology in Russian. The output from an MT system is unlikely to be very good, but for technical readers who know enough about the field, who know what is going on generally in this science, and who can maybe even guess roughly what the article is about, it may well provide sufficient material to get at least some idea of the content of the text.

The first empirical study to identify the importance of the user's familiarity with the textual context of a raw machine-translated text was Henisz-Dostert's 1979 survey of scientists, which inquired, "What does understandability of MT primarily depend on?" Participants could choose from language-related choices (sentence structure, translation of words, and general style), a layout-related choice (format), and the choice of *familiarity with subject matter*. The researcher had a clearly stated goal to prompt participants to evaluate the importance of linguistic versus extralinguistic factors. A total of 88.5% of participants ranked *familiarity with subject matter* as the most important factor determining understandability of machine-translated texts,

[29] For example, Nord (1991, 53) included background knowledge of recipients as part of text analysis, DuBay (2004, 28) cited prior knowledge as a contributor to text readability, and Moravcsik and Kintsch (1993) discussed the role of domain knowledge in text comprehension.

prompting Henisz-Dostert to comment that, "this should come as no surprise, not because machine-translated texts are involved here, but because much of our understanding of any phenomenon, and especially linguistic ones, depends on context" (Henisz-Dostert 1979, 188).

Subsequent studies also highlighted the effects of familiarity with textual context on MT reception. In a 2003 study on users of an MT service at PricewaterhouseCoopers, Smith (2003) concluded that the reaction of first-time users of MT depends on several factors, one of which is the source document's subject matter. Bowker and Buitrago Ciro's 2019 book on the use of MT by academic researchers reported that researchers adopt MT for understanding scientific articles in languages they do not know well. Bowker and Buitrago Ciro (2019, 80) stated that:

> Moreover, because they are domain experts who are already familiar with the concepts in their field, the resulting machine translation output is typically seen as being quite usable and helpful for comprehension purposes. Therefore, when viewed through the lens of information assimilation, machine translation is often seen in a quite positive light.

The results described here raise an issue about the potential use of raw MT in health care scenarios. As discussed in Section 5.4, a common theme in the literature on the possible use of raw MT in this area is that the risk of mistranslation and the ensuing consequences are so high that raw MT is often seen as inappropriate. If, as the results in this section indicate, familiarity with textual context promotes the successful use of raw MT, then a second important risk factor involved in the use of raw MT in health care would be a target audience's lack of contextual knowledge. In other words, if patients and other targeted readers of information on health care are unfamiliar with genres such as patient instructions and subject matters such as specific diseases (which undoubtedly at least some readers are) their ability to successfully access raw machine-translated material can be reduced.

6.1.3 Users had a good level of MT literacy

A third user quality that was observed or reported to influence the use and reception of raw MT was MT literacy. In short, MT literacy refers to knowledge about how MT works, how it can be used in a particular context, and what the implications are of using it for various purposes (O'Brien and Ehrenberger-Dow 2020, 145). For a full definition and discussion of the concept, see Section 2.3.2.

In the study that led to Article II of this dissertation, the patent professionals who acted as informants were observed to have a relatively high level of MT literacy, which influenced their ability to use raw MT effectively. The informants displayed an understanding of the basics of MT technology, were adept at accessing different MT tools, and were aware of the increasing improvement in the technology. Some mentioned noticing improvements themselves. Another important display of MT literacy was an awareness that MT can produce errors and an ability to spot potential errors. In the following example, informant PP1 described their reaction to a translation from Google Translate:

> Now that can't be, could it really work like that? When you know something about the technology [the technology described in the patent document], it doesn't work like that. It doesn't go like that. And then you go to the European Patent Office's [MT tool] and you see that, OK, yep, translation error."

Besides observing that a high level of MT literacy proved useful in individual informants' use of raw MT, Article II also identified a collective MT literacy across stakeholders in the IPR ecosystem as an influencing factor in MT gisting. This is further examined in Section 6.3.2.

6.1.3.1 Related work

Non-empirical research on MT literacy has focused on defining the concept and on outlining potential applications and benefits. A full discussion on definitions can be found in Section 2.3.2.

Tinsley (2017, 410) discussed how MT education could promote a more "mainstream acceptance of MT" and how part of that should involve identifying suitable contexts for MT gisting, such as the patenting context. Martindale (2020) identified the training of MT users as a way to mitigate the risks that are involved in MT gisting. Two recent studies proposed new contexts in which MT literacy would be needed. Vieira, O'Hagan and O'Sullivan (2020) observed that the concept of MT literacy as coined and defined for academic contexts by Bowker and Buitrago Ciro (2019) was useful, yet the limitation of its scope to educational contexts left out a large number of people who would benefit from a better understanding of MT. They highlighted the need for a similar promotion of MT literacy for medical, legal and other high-stake settings. In a similar vein, O'Brien and Ehrensberger-Dow (2020) suggested that MT literacy training would benefit raw MT users such as people involved in crisis communications and patenting scenarios.

Several empirical studies found evidence that increased experience and knowledge of MT influenced users' reactions and ability to use MT. Henisz-Dostert's (1979) survey of scientists included several questions that aimed to establish whether experience with MT would lead to an improved ability to use machine-translated information. The survey found that a small majority of respondents had noticed improvements in MT quality over the time they had used it. A larger majority, 76%, reported that they found it easier to read the later translations they had received than the first ones. Finally, 97% reported that they felt they could get used to the style of MT (ibid., 202). This evidence suggested that the ability to assimilate information from raw MT output influences reception positively, and that the ability can also grow with increased experience. In a second paper, Smith (2003) noted that first-time MT users at PricewaterhouseCoopers had varying reactions to raw MT depending on a number of factors, including "users' initial expectations based on the extent of their knowledge of similar systems and their faith in IT in general" (Smith 2003, 13). Finally, Yasouka and Björn (2011) reported that students involved in an experiment with MT-mediated communication quickly learned to adapt the language in their messages to get better MT output and that this contributed to successful communication via MT.

Other researchers have highlighted how MT tool manufacturers' efforts to train their users about MT has contributed to the successful use of those tools. For example, Flournoy and Callison-Burch (2000) outlined factors that contributed to the success of the Amikai system, including the fact that "the system trains users to understand better the strengths and limitations of the MT engines" (ibid., 2). In a similar vein, Gaspari (2004) indicated that Babelfish had a noticeable advantage in usability over contemporary MT systems because it offered users tips on how to use the system in an optimal way.

6.2 User goals and tasks

The second category of context of use provided by ISO 9241-210:2019 are the goals and tasks of the user, referring to "the way in which users typically carry out tasks, the frequency and duration of performance, interdependencies and activities to be carried out in parallel" (ISO 2019, 13). In this dissertation, the category was manifested primarily in the process-oriented parts of this description, namely, in the way users carried out tasks and in the auxiliary activities they adopted in parallel to their use of MT. Several studies in this dissertation revealed auxiliary activities that

users employed as strategies to help them achieve an understanding of texts. Future studies may reveal

An important aspect of this category is that it applies to both the process of using MT and the higher-level tasks a person is engaged in when they employ MT gisting. People rarely use raw MT as a stand-alone task; they access MT when they are involved in another process in which information in another language is needed. The higher-level processes involving information that needs to be machine translated vary widely. Examples from this thesis included processes such as reading academic articles that are published in other languages, assessing the relevance of patent documents in other languages, and accessing important health and safety information that is in a language the user does not fully command.

The contextual influences related to users' goals and tasks are listed in Table 7.

Table 7. Contextual influences related to user goals and tasks

Factor	Section in which factor discussed
Users accessed multimodal information in source texts	Section 6.2.1
Users accessed outputs from different MT tools	Section 6.2.2
Users adapted source texts to get better MT output	Section 6.2.3
Users negotiated meaning during the process of using MT	Section 6.2.4

6.2.1 Users accessed multimodal information in source texts

The first task-based contextual influence uncovered in this dissertation involved the parallel activity of accessing multimodal elements of source documents to augment understanding of machine-translated texts. Such information included pictures, formulas, and equations, which can be understood even without competence in the source language. This phenomenon was reported in Articles II and III of this dissertation and has been discussed in various past studies as well.

In Articles II and III, which both drew on the same data set of interviews with nine patent professionals, informants described how accessing the multimodal elements of source documents was one of the primary tools they use to augment their understanding of the MT output of patent documents. They reported relying on original patents' illustrations, as in this example from PP6 (Nurminen 2019, 37):

> When it's good enough that I can see that it's relevant? It's a combination of understanding the figures and understanding the machine translated text.

A second multimodal element patent professionals accessed were mathematical and chemical formulas. This was illustrated by informant PP7 as follows (Nurminen 2020, 110):

> It's a combination of the original, if I see the chemical formulas or whatever they are using, because the abbreviations, they are not translated, like in carbohydrate variations, they are not translating those […] and then I have the original and I have the translation. So then I combine them.

6.2.1.1 Related work

Several non-empirical studies discussed the benefits raw MT users might get from accessing non-verbal elements in the source text to augment the information they receive from the machine-translated texts. For example, List (2012, 195) argued that multimodal information can be useful when reading machine-translated patent documents:

> However important understanding the text of the document is, it should also be remembered that there are always the non-text elements to consider – the drawings, formulae, mathematical symbols, etc. – these are universally understood and should not be overlooked as important tools for reviewing [patent] documents in any language.

In the same vein, Way (2013) proposed that raw MT could be sufficient for translating product descriptions in eBay because they are very often accompanied by photos of the products, implying that information that might be lacking due to poor MT output would be compensated by multimodal elements of description. Gao et al. (2015, 861) suggested that MT-mediated communication might be improved through multiple channels of communication such as "images that illustrate some part of the original (untranslated) message along with the translations to help with message interpretation."

A few empirical studies uncovered evidence that multimodal elements in source texts could enhance people's understanding of raw MT. Henisz-Dostert (1979) asked scientists who use MT gisting if the understandability of MT was most affected by untranslated words, incorrectly translated words, lack of multimodal elements such as formulas and figures, or sentence structure. The study found that, although the multimodal elements were not as important in users' opinions as the other factors, a small percentage of respondents did indicate that the lack of those elements would be the most or second-most important factor in understandability. Suzuki and

Hishiyama (2016) found that replacing some words with numbers and symbols contributed to successful communication via MT. Parra et al. (2018) developed a hand-held translation system which aimed to help people manage special diets while traveling. Besides MT, the system included images and ingredient lists to help ensure accuracy and understanding. In a user test, participants were tasked with identifying items on a menu that might contain ingredients that are sensitive for special diets. The translation system with images and ingredient lists was tested against the same system's MT without the extra information and against Google Translate. The results revealed that the system with MT, images and ingredients lists was the most effective.

6.2.2 Users accessed outputs from different MT tools

A second phenomenon related to users' goals and tasks, again related to parallel activities undertaken while using raw MT, was the practice of entering a source text into two different MT tools and using the multiple outputs to arrive at a meaning of the original text. This practice was reported in Articles II and III as well as in past research, and the benefit of showing two outputs to raw MT users has also been studied empirically. Researchers have additionally contemplated the cognitive processes involved in comparing alternate translations, but the topic has not been empirically studied to date.

In the study that led to Articles II and III, the patent professionals who acted as informants reported that they occasionally translated texts with different MT tools and used the multiple outputs to arrive at a meaning of the source text. The MT tools the informants most often worked with have been developed specifically for translating patent documents and they are most often embedded in the main tool patent professionals work with, patent databases. However, a common theme in the study's interviews was the practice of translating the same text with a second MT tool to augment understanding. The most common secondary tool informants reported using were MT tools provided by individual governments' own patenting offices, as reported by informant PP6 (Nurminen 2019, 37):

> …for instance if it's a Chinese document I go to Chinese Patent Office website and try to find the same application there…and usually it's a different machine translation and that actually helps sometimes; when you have two machine translations you can read them at the same time and maybe it gives you a better impression.

Recognizing this as a common strategy, at least one patent-database manufacturer provides a link to a second MT engine, Google Translate, directly in their own MT tool's user interface (Nurminen 2020, 113).

6.2.2.1 Related work

In a presentation on the responsible use of raw MT, Martindale (2020) proposed that one technique for mitigating the risk involved in using raw MT would be a "self-validation" of MT output before taking any action based on the machine-translated information. One of the means suggested for self-validation would be eliciting output from multiple MT tools, and Martindale proposed that interfaces that show multiple outputs should be used to encourage users to do this. She also provided anecdotal evidence that users prefer to see multiple outputs side-by-side.

Empirical studies have both observed a tendency among MT users to access multiple MT outputs and tested the effect of displaying two outputs to users. Similarly to Articles II and III, Tinsley et al. (2012) reported a tendency for patent professionals to access multiple MT outputs when reviewing patent documents, while Anazawa, Ishikawa and Takahiro (2013) noted the same tendency among nursing professionals in Japan.

Other researchers have empirically tested the effects of showing users two MT outputs instead of one. Xu et al. (2014) and Gao et al. (2015) conducted experiments in which people engaged in MT-mediate communication were shown either one or two translations of their conversation partners' messages. Xu et al. (2014) found that users reacted positively to being shown two alternative translations and that the use of two translations did not add to the time required for completing a conversation. Gao et al. (2015) reported that two-person teams who were shown two MT outputs performed better than teams using only one output, and they did not experience a higher workload from working with two outputs instead of one.

All of the studies listed in this section have also explored possible explanations for what actually occurs when people use multiple MT outputs and how they derive meaning from the alternative translations. Tinsley et al. (2012, 3) suggested that users were acquiring a "second opinion", though did not detail how they might be comparing the two translations and deciding which is correct. Articles III and IV of this dissertation (Nurminen 2019; 2020), Xu et al. (2014) and Anazawa, Ishikawa and Takahiro (2013) all proposed that users were either reviewing multiple outputs and selecting the best one, or they were combining pieces they understood from different outputs and constructing meaning from the combination. Xu et al. (2014) and Gao

et al. (2015) suggested that the parts of translations in which two MT outputs were similar might increase users' confidence in those parts, while in places where neither translation is understandable, users would benefit from the clear indication that there is something wrong. Martindale (2020) likewise proposed that the identification of passages in which multiple MT outputs result in different translations would help raw MT users to pinpoint areas that require further investigation. These proposed explanations highlight our lack of knowledge on the cognitive processes involved in MT gisting with multiple outputs and indicate a need for further study.

6.2.3 Users adapted source texts to get better MT output

It is generally agreed that the quality of the source text affects the quality of the MT output. It would logically follow that adapting source texts so that they are more suitable for MT would lead to improved MT output. One line of MT research has focused on formal efforts to adapt source texts, such as applying controlled language to static texts, to see if they lead to improvements in output quality. More informal efforts to adapt language for improved MT output have been detected in research on MT users, including in Article IV as well as in past studies. This group of studies has noted a tendency among people engaged in MT-mediated communication to adapt the language in their own messages in the hopes of attaining machine-translated output that is easier for their counterparts to understand. Past research has also illustrated some of the types of adaptations people make to language and explored technology that facilitates adaptation.

Article IV of this dissertation detected a tendency for participants in MT-mediated interviews to change messages to facilitate understanding or to repair communication. The article also uncovered evidence that participants learned to use the strategy over the course of the interviews and that the adaptations contributed to successful communication. A few examples from Article IV illustrate the influence of these factors on the communication in the interviews.

In the first example, I asked an informant how they felt MT had worked in the interview. When the informant replied that he did not understand the question, I pinpointed the word *feel* and its translation as *se siente* as the potential problem. I therefore adapted my language, repeating the question but using the word *think* instead of *feel*. After the adaptation and translation as *cree*, the informant was able to answer the question, an indication that the adaptation successfully aided understanding.

In the second example, an informant (Tomás, a pseudonym) discussed a three-page document he did not think was translated well, using the word *hojas* in Spanish, which was translated into *leaves* in English. I was nevertheless able to guess the correct meaning and I replied using the English word *pages*, which the system translated into *páginas*. Ten minutes later in the interview, Tomás again used the Spanish word *hojas*, then corrected himself to attain a translation that would match the word I used. This is illustrated in Figure 7. below. In this excerpt, informant Tomás's input, in Spanish, is shown on top. Below that is the translation into English.

Figure 7. Informant Tomás uses *hojas* again but corrects himself with *paginas*. Bolding added by author to highlight words involved in adaptation

Tomás
9:01 PM

si la descargue, traduje un paper de 3 **hojas**
If the download, I translated a paper of 3 **leaves**

perdon, 3 **paginas**
Sorry, 3 **pages**

A further 20 minutes into the interview, Tomás used the word *página*, proactively adapting his language to produce MT output that would be more understandable to the interviewer. These examples suggest that competence in the language of the other participant influences communication and enables the use of adaptation to improve translations.

6.2.3.1 Related work

One component of the definition of MT literacy proposed by Bowker and Buitrago Ciro (2019, 88) was the ability to "create or modify a scholarly text so that it could be translated more easily by a machine translation system". The researchers saw a need for heightened awareness of the need to produce more user- and machine-friendly academic texts to make them more accessible to scholars who do not have a high proficiency in English. Bowker and Buitrago Ciro also presented guidelines for the types of adaptations authors should make to achieve this. A few manufacturers of commercial MT solutions, such as Amikai, have encouraged their

users to learn adaptation techniques. As reported by Flournoy and Callison-Burch (2000, 4–6), the Amikai system had built-in functions for encouraging users to learn the benefits of adapting their language for better output and for offering feedback to them on their language. The AltaVista translation system's FAQ page encouraged users to "use short sentences and avoid slang, idiomatic expressions, and unnecessary synonyms" (Yang and Lange 2003, 200).

In empirical research, a study by Nomura, Ishida and Yasouka (2003) that belonged to the ICE2002 project (for more on this project, see Section 3.1) noted that participants in MT-mediated conversations made changes to the language in their messages so that they would be more easily translated by the MT software. The article also described an example modification, in which a participant added subjects to their Japanese sentences although it would be natural to omit them in Japanese.

Several subsequent articles discussed specific adaptations MT users made to their own messages. In a second ICE2002 study, Ogura et al. (2004), performed a specific examination of the types of adaptations participants made and found that these adaptations differed between the speakers of different languages. In his recollections of the ICE2002 project, Ishida (2016, 6) described how Japanese participants started using far more personal pronouns than are normally used in Japanese. Yasouka and Bjorn (2011, 113) reported on a strategy of simplifying and shortening messages in their experiments. Similarly, Suzuki and Hishiyama (2016) illustrated how a study participant who was imparting expert knowledge via MT adapted their language for the MT by trying to use simple sentences and breaking up sentences when one sentence would hold a great amount of information. This strategy was also found to aid communication in the experiment.

In some of the studies mentioned in this section, the participants in MT-mediated communication who adapted their messages had some knowledge of their conversation partner's language. This helped them to detect problems and evaluate whether their adaptations were successful. However, participants do not always have language competences to help them, and attempts have been made to provide technology that would make it possible for MT users to evaluate the translations of their own messages even when they have no understanding of the other language. In the first example, again taken from the ICE2002 project and reported in Ogura et al. (2004), participants did not know the languages their messages would be translated into. However, many had some level of competence in English, which was the pivot language used by the MT system, so the system showed them the pivot language translation. They could then adapt their own messages until they were

satisfied with the English translation, after which the system translated the message to the target language.

The second method employed to give participants in MT-mediated communication a way to evaluate how their messages were being translated was to include a back-translation function. Shigenobu (2007) and Hautasaari (2010) implemented such solutions in the tools used for their studies. After users wrote a message in their own language, the system would translate it, then perform a back translation into the original language so that the user could confirm if the message was accurate. If it was not, they could adapt their message and try again until they were satisfied. Shinegobu (2007) found that the improvements gained through back-translation depended on the language pair involved, while Hautasaari commented that the feature helped "make communication fluent and natural" (Hautasaari 2010, 76).

6.2.4 Users negotiated meaning during the process of using MT

A final task-related contextual influence that this dissertation revealed was the practice of negotiating the meaning of machine-translation information among multiple people while using raw MT. The practice was detected in Article IV, which focused on MT-mediated communication. Past research has also discussed the practice specifically in the dynamic context of MT-mediated communication. However, Articles II and III of this dissertation uncovered a practice of negotiating for meaning with more formal and static texts in MT for assimilation. To the best of my knowledge, no other study has discussed this phenomenon in relation to MT for assimilation.

Research on second-language acquisition defines *negotiation for meaning* as "a series of activities conducted by addressor and addressee to make themselves understand and be understood by their interlocutors" (Yufrizal 2001, 63). The phenomenon has been studied widely in the areas of second-language acquisition and native/non-native speaker interaction. Long (1983) defined several tactics that tend to be used in native/non-native conversations to negotiate meaning, including (ibid., 137–138):

> Requests for clarification: expressions, often questions, designed to elicit clarification of preceding utterance
>
> Confirmation of understanding: expressions designed to elicit confirmation that a preceding utterance has been correctly understood
>
> Repetition: repetition of part or all of conversant's preceding messages, either exactly or paraphrased

The interviews in Article IV revealed evidence of these same tactics being used in MT-mediated communication. Both the interviewer and the informants displayed attempts to understand, and to ensure they understood, through asking questions, repeating and rephrasing, and requesting clarification. In one example, an informant uses the term *SLA* in a message to me, the interviewer. I rephrase and ask for clarification with the question "service level agreement?" and get the response that I am correct. The informant then describes what they will do with the SLA: "It is a document that provided me with, I'm doing my own paper and needed a guide." I again rephrase and ask for confirmation that I've understood by asking, "So you needed information on how to create an SLA?" (Nurminen 2016, 76). The article also noted that although these tactics resolved a number of issues with gaps in understanding, the participants did not manage to resolve all issues and some questions remained after the interviews were concluded.

Meaning-making through negotiation was also a theme uncovered in the research on patent professionals conducted for this dissertation. Article II described how the meaning of raw machine-translated patent documents was discussed, examined, and sometimes challenged during IPR processes. Article III centered on the theme of negotiation for meaning, exploring MT gisting in the IPR ecosystem through the lens of the concept of distributed cognition. It was, in fact, the discovery of the concept and its potential applicability to the patent case that led me to analyze the concept of meaning-making more deeply. The article identified the human actors patent professionals include in negotiations on the meaning of machine-translated texts. One such group are the technical experts behind the inventions being patented or whose patent needed to be protected. Study informants described how they often arrived at an understanding of texts through discussions with these experts.

The second group of people patent professionals negotiated meaning with were people in the larger network of stakeholders involved in the IPR case they were working on. The patenting process often involves a large network of IPR service providers as well as national patenting offices in various countries, and patent professionals could negotiate with people in that network when trying to understand

a patent document. Patent professionals even negotiated meaning with their competitors. The patenting process allows competitors to propose a different interpretation of relevant documents or challenge the interpretation of other parties' documents. This occurs for both English-language as well as translated and machine-translated documents. The ensuing discussions can help formulate the interpretations of the meaning of texts, including machine-translated ones. The concept of distributed cognition proved to be a fruitful avenue for exploring how meaning is negotiated in the IPR ecosystem.

In this section I have focused on how negotiation for meaning was found to be an influencing factor within raw MT users' own MT processes. However, the fact that negotiation for meaning was common in the environment in which MT gisting took place was also found to influence the use of raw MT. This is discussed in more detail in Section 6.3.3.

6.2.4.1 Related work

In the literature on MT for communication, researchers have studied negotiation for meaning through conversants' use of repetition, rephrasing, and requests for confirmation of understanding, and also through the concept of common ground. Originating in human communication theory, the concept of common ground was adopted in research on electronic communication (Monk 2009) and MT-mediated communication (for example, Yamashita and Ishida 2006a; Yamashita et al. 2009). Common ground is defined by Clark (1993, 93) as, "Two people's common ground is, in effect, the sum of their mutual, common, or joint knowledge, beliefs, and suppositions." Common ground is maintained and updated throughout a conversation through the act of *grounding*, in which participants trying to assure themselves that their messages have been understood. This ensures that not only are the participants aware that their own understanding of the knowledge and suppositions is shared with other conversants, but that they are also aware that the other conversants are aware of this joint knowledge.

In Nomura, Ishida and Yasouka 2003, one of the first studies in MT for communication, first noted a tendency among participants in MT-mediated conversations to employ various strategies for correcting translation errors. One such strategy was asking for confirmation of understanding. Calefato, Lanubile and Prikladnicki (2011) and Calefato et al. (2012) compared the MT-mediated conversations with conversations held in the lingua franca of English. They examined evidence of a lack of common ground through what they termed

clarification dialogues, which are "typically initiated by [conversants] rephrasing statements in their own words, often in form of questions" (Calefato, Lanubile and Prikladnicki 2011, 99). They hypothesized that MT-mediated conversations would lead to an increased number of clarification requests. Their results, however, were inconclusive.

One group of studies applied the concept of common ground as a method for evaluating success in MT-mediated conversations. Yamashita and Ishida (2006a) investigated grounding through the use of referential expressions, in discussions conducted in English and those conducted via MT. They focused on the language used to refer to specific items in the conversation, whether the referential expressions were understood and adopted by the other conversants, and whether conversants would display a tendency to shorten referential expressions, as they do in human communication. They found that this method of grounding was more difficult in MT-mediated conversations, since participants were not working with fellow conversants' utterances but with translations of those utterances, and those could not be reproduced exactly nor could they easily be shortened. Yamashita and Ishida also hypothesized that conversants in MT-mediated conversations would have more questions and requests for confirmation than people conversing in English, but this hypothesis was not confirmed.

Yamashita et al. (2009) applied similar methods to conversations involving multiple participants and multiple languages and found that establishing and maintaining common ground was even more difficult than in conversations with only two participants. In an experiment with 15 teams consisting of Japanese-speaking students and English-speaking counterparts who were given a task that required them to communicate via MT, Yasouka and Björn (2011) noted that, although all teams mentioned communication difficulties due to MT, 14 out of 15 managed to complete the task. They concluded that having a shared task and the development of shared jargon provided common ground that helped them complete the task, and that "communication relies more on a dynamic process where participants establish common ground than on reproducability and grammatical accuracy" (ibid., 110).

Gao et al. (2015) searched for positive and negative evidence of grounding in conversations conducted in three different ways: in a lingua franca (English), via MT in a tool which shows the conversant the output of one MT tool, and via MT in a tool that shows the conversants outputs from two different MT tools. They considered acknowledgements of understanding to be positive evidence of grounding, while questions and requests for confirmation would be negative

evidence of grounding. Their results showed that conversations involving two different MT outputs did result in more positive evidence of grounding than when only one output was shown, though the difference in positive evidence was similar for conversations with two outputs and those in English. The results for negative evidence, as displayed through questions and requests for confirmation, were mixed.

As is evidenced in this short review of research, although there is consensus that negotiation for meaning occurs with raw MT, mixed results have been obtained on how, exactly, negotiation for meaning influences MT-mediated conversations. Further study is needed.

6.3 Environmental factors

The final category of context of use outlined in ISO 9241-210:2019 is environment, which includes the technical, physical, and social and cultural environments. The technical environment is comprised of the hardware, software and other materials accessed while working. In this dissertation, that refers to not only the MT tools themselves, but also to the higher-level software or platforms that the MT tools might be embedded in, additional features offered by the MT tools, and other tools that users rely on while they are engaging with MT. The social and cultural environment, defined by ISO as "factors such as work practices, organizational structure, and attitudes" (ISO 2019, 31) was also important in influencing the use of raw MT. Aspects of physical environments that would have influenced MT gisting were not detected in this dissertation and are therefore not included here.

The factors related to the environment that were found to influence the use and reception of raw MT are listed in Table 8. below.

Table 8. Factors in the environment that were found to influence raw MT reception

Factor	Section in which factor discussed
Users benefited from auxiliary technology available in their environment	Section 6.3.1
Legitimacy of the practice of MT gisting in the environment	Section 6.3.2
Raw MT was used in an environment in which negotiation for meaning is a common practice	Section 6.3.3
Raw MT was used in a risk-tolerant environment	Section 6.3.4

6.3.1 Users benefited from auxiliary technology available in their environment

While the technical details of MT engines already have a large body of literature devoted to them, in this dissertation I was interested in the other features and tools which were available in the context of use and which affected how people engaged with MT. A number of such auxiliary features and technology were observed or reported to influence the use and reception of MT in the dissertation research. This included the technical environments in which the MT tools were embedded, additional features of MT tools, and auxiliary technologies that informants used while they were engaging with raw MT.

Related empirical studies have also been conducted to test the effects of auxiliary aids such as functions that highlight keywords or pieces of text that were problematic for MT or display outputs from two different MT tools. The studies focused on a wide variety of auxiliary tools and at this point, it is difficult to draw collective conclusions on any one feature. However, as the quality of MT continues to improve, it becomes more difficult to achieve noticeable improvements in the user experience through linguistic quality alone. Augmentations in the surrounding technology might well bring benefits that prove to be immediately helpful to end users, and many of these might require little effort to implement. For these reasons, further investigation into the influence of these auxiliary technologies is needed.

The first aspect of technology that was found to influence MT gisting in this dissertation was the availability or integration of MT tools in the upper-level workflows in which raw MT is used. Both Articles II and III discussed how the manufacturers of patent databases, the main tools used by patent professionals in their work, have facilitated the integration of MT into IPR workflows by providing

MT directly in their tools. Study data provided ample descriptions of how MT is indeed tightly integrated into patent professionals' everyday processes and how that integration makes working with raw MT seamless within their higher-level IPR processes. The integration of MT into the chat tools used for interviewing was also important to the interview context that was studied in Article IV. The tool chosen for interviewing was Skype Translator because it was a commonly available tool. The embedding of MT directly into the chat tool also meant that it was easy to use, even for people inexperienced with MT.

Various other auxiliary tools were reported to enhance patent professionals' use of raw MT. One patent database manufacturer, recognizing that patent professionals often elicit translations from a second MT tool, integrated access to a second tool directly into their own MT user interface. A second auxiliary tool in that environment allowed patent professionals to save and share machine-translated patent documents among various people in the patent context, including colleagues, customers, and patent applicants. To ensure that the eventual readers of such documents are aware they are reading raw MT, such documents are clearly marked as being MT with "colored frames, stamps labeled [machine translation], and documents with both original texts and translations shown in consecutive paragraphs or side-by-side columns" (Nurminen 2020, 108). Often machine-translated documents also contain a stamp with the name of the system that was used for translation and a time stamp. These can be important indicators to patent professionals, whose awareness of the increasing quality of MT systems leads them to occasionally re-translate documents that have older time stamps.

6.3.1.1 Related work

A number of studies have described the auxiliary features or technologies included in MT systems. The ICE2002 project's technology, which machine-translated between different Asian languages, had a feature that showed users the results of the translation from their own language into the pivot language of English. Study participants, most of whom had at least a basic understanding of English, could review the English translations and adapt their messages until they were satisfied, and only then would the tool translate from English into the target Asian languages (Nomura, Ishida and Ysaouka 2003; Ogura et al. 2004). Studies by Shigenobu (2007) and Hautasaari (2010) employed technology that allowed users to back-translate their messages into their own languages so that they could evaluate whether their original messages led to understandable translations. If not, they could edit the original

messages and re-translate to get better MT output (for more, see Section 6.2.3). Tinsley et al. (2012) described a tool for patent translation that included a feature that highlighted segments of the source and target texts that corresponded to each other. Highlighting has also been used to show parts of the source text that caused translation problems (Shigenobu 2007) and to show keywords (Gao et al. 2013).

Empirical studies have also focused on testing auxiliary technology. Two studies (Xu et al. 2014; Gao et al. 2015) tested and confirmed that showing raw MT users the output from two different MT systems was more beneficial than showing them only one output. Yamashita and Ishida (2006b) studied instances of misconception in MT-mediated conversations and proposed a method for automatically predicting places where misconception can occur in MT-mediated conversations. Finally, Pituxcoosuvarn et al. (2018) investigated what children who were using MT to communicate with other children turned to when the MT failed. Based on their observations, they proposed various auxiliary tools to help communication. Among those were a shared image browser so that conversants can search for and share pictures, and a feature that would show MT output in both the user's own language and a second language they might know.

6.3.2 Legitimacy of the practice of MT gisting in the environment

The second environment-oriented contextual influence revealed in this disseration, which I termed *legitimacy*, was discussed in Articles II and V. As described in Section 5.2.1.1, the general definition of legitimacy involves the concepts of allowability or acceptability of a practice and an adherence to agreed-upon rules when conducting the practice. Another aspect is that the legitimacy is ordained by and constrained to a specific group. In Article II, I developed a more specific definition of legitimacy for the context of patent professionals' work. In it, legitimacy was comprised of transparency, which stems from and also contributes to acceptability; generally agreed upon boundaries of legitimacy, which refers to the constraint of legitimacy to a group; and collective MT literacy, which is fostered through guidelines and training. In Article V, I examined legitimacy through the concept of acceptability. At this point in my research, legitimacy remains a collection of related concepts, with one or more concepts being applicable in different contexts of MT gisting. However, despite its imprecise nature, legitimacy proved to be a useful concept in analyzing MT gisting in this dissertation, providing a nuanced way of describing individual and collective outlooks on MT gisting.

One of the conclusions of Article II was that practice of MT gisting enjoyed a high level of legitimacy in the IPR environment. As described in Article II, this legitimacy was one of the factors that gave affordances for the use of raw MT, and such affordances appeard to play an important role in making the use of raw MT tenable (Nurminen 2019, 38). A further discussion of legitimacy in this environment can be found in Section 5.2.1.1 and in Article II.

Article V discussed legitimacy through the concept of acceptability. The article, which explored the use of MT for increasing accessibility to information for groups who are underserved, examined acceptability as an ethical consideration when using MT. People and projects involved in efforts to translate more information for a group of people should recognize that these efforts need to be acceptable to that target group, and that acceptability can be affected by various factors. One of the most important influencing factors discussed in the article was the reason target audiences have for wanting information in their own language: people who want information translated as a way to preserve their cultural heritage were less likely to find the use of MT acceptable than people who want translations to access information they need (see Bowker 2009). A second, closely related, aspect of acceptability was perception. Offering information in raw MT for a certain language might lead to a perception that the language has a lesser status than languages for which information is translated by humans, which can affect the acceptability of MT for speakers of that language. A final aspect of acceptability examined in the article concerned the application of MT to specific subject matters such as public health and medicine. The article examined how the subject matter of the material translated via MT also affected acceptability. Information on topics such as public health and medicine were cited as being so risky that the practice of using MT was seen as unacceptable.

One important note about the acceptability of a practice as examined in this dissertation is that it differs from the acceptability of MT that has been studied in research such as Castilho (2016), Castilho and O'Brien (2017), and some of the studies mentioned in Birch et al. (2018). Those studies focused on the acceptability of MT output, or specific texts, as measured through scores given to the texts for usability, quality and satisfaction. In the present dissertation, I examine the acceptability of the entire practice of using raw MT and not the acceptability of specific machine-translated texts.

6.3.2.1 Related work

Past MT research has explored the legitimacy-related concepts of transparency and acceptability. Each of these is examined separately below.

The need for MT to be transparent has been discussed throughout its history. Yang and Lange (1998; 2003) reported on suggestions by the users of one of the first freely available online MT tools, AltaVista Translation with Systran, that machine translations should be marked with a warning that the text was machine-translated. As reported by Morland (2002), one of the ground rules for the NCR Corporation's implementation of MT for disseminating its internal newsletter was that users would be aware that they were receiving raw MT. Some have suggested that even post-edited MT should be equipped with a note that the information was produced with MT (Bowker and Eghoetz 2007, 220). DePalma (2007, 50) cited attempts at transparency that were problematic, such as presenting users with a note that content is machine-translated, but only providing the note in English.

Transparency has also been described as having a role in the use of MT in the discovery process in legal settings (for more on this process, see Section 3.3.3). This emerging use case for raw MT is similar to MT gisting the IPR ecosystem. Both involve using raw MT to sift quickly through large amounts of information in order to pinpoint relevant parts, which might then be used in their machine-translated state or might be sent for human translation. The practice of using raw MT in this ecosystem is growing and, as described by John Tinsley (Beninatto and Stevens 2019, 9), it is fully transparent:

> And so, the fact that you're using MT is fully transparent. So, you can go into the court and say to the judge, "We are taking this position on the basis of a machine translation of this document into English," and that's legally defensible.

The effects of transparency on raw MT reception have also been tested empirically. In Gao et al. (2014), pairs of participants communicated via a chat application, with each participant either writing messages in the lingua franca of English or in their own native language, in which case the message was then machine-translated into English. The participants were given differing information on whether their communication was all in English or whether the other person's communication came through MT. The results of the study showed that the belief that MT was involved led speakers to attribute miscommunication less often to their conversation partner than if they did not believe MT was involved. This belief also resulted in a generally more positive communication experience. Rosetti, O'Brien and Cadwell (2020) tested how the awareness of raw MT or post-editing affected participants'

trust in and comprehension of crisis-related messages. The participants reviewed three versions of a message with instructions for action in a crisis scenario. One of the messages was in English while two were machine-translated into Italian and then post-edited. However, participants were told that one of the post-edited versions was actually raw MT. They were then asked to indicate how comprehensible each message was, how much they trusted it, and how likely they were to follow the instructions. The study found non-significant differences in trust and comprehensibility between the three versions of the messages.

Acceptability has also been examined in past research on MT and other technologies, for example, in Nielsen's discussion of *social acceptability* (Nielsen 1993, 24–25) and in Suojanen, Koskinen and Tuominen's (2015, 16) subsequent claim that in translation, social acceptability takes on a more important role because of its intercultural nature. "Even if a product is highly usable, it may lose its utility completely in a target culture if its use is not socially acceptable."

A number of empirical studies have also focused on the acceptability of the use of MT. One group of studies (Bowker and Eghoetz 2007; Bowker 2009; Bowker and Buitrago Ciro 2015) compared different groups' views on the acceptability of different forms of MT and HT. The studies applied the idea of a *recipient evaluation,* in which participants are given different versions of a text (human translated, maximally post-edited, rapidly post-edited and raw MT), as well as information on the cost and speed used to produce each version. Weighing all of these inputs, participants were asked to evaluate which version of the text would best meet their needs and for what reasons. The research found that, in addition to linguistic quality, acceptability was influenced by the genre of the texts used in the experiments, whether the participants considered themselves to be language professionals or not, and the reasons participants wanted the texts in question translated. In a second group of studies in the HimL project[30], the acceptability of MT use in the public health domain was examined. Although machine-translated information in one study was judged to be generally useful in three out of four languages, a second study concluded that MT errors and their inherent risk render MT an unacceptable method for publishing public health information (Birch et al. 2018, 27).

Besides the individual concepts of transparency and acceptability, one recent paper cited several of the same elements of legitimacy that were defined for Article II. In the article, Vieira, O'Hagan and O'Sullivan (2020) examined the use of MT in medical and legal settings and suggested the need for improvements in MT literacy,

[30] Health in my Language project: http://www.himl.eu/

official guidelines, and "robust standards regarding the situations in which MT use is and is not admissible" (ibid., 13) in those environments.

6.3.3 Raw MT was used in an environment in which negotiation for meaning is a common practice

Section 6.2.4 examined how users of raw MT negotiated with others about the meaning of machine-translated information. However, Article III concluded that a second factor that influenced MT gisting in the IPR context was that negotiation for meaning occurred in an environment in which negotiation for meaning is a common and accepted practice. In IPR processes, even when all information is in English, its interpretation is proferred, debated, challenged and changed. When that information is accessed through raw MT, it adds some complexity to the negotiations, but the base nature of the discussions remains and is an accepted part of the patenting process. In fact, the process has built-in phases for analysis of information, and that provides space for raw MT to be examined and potential errors to be spotted. Participants in this process are also accustomed to achieving meaning through negotiation, which offers further support for the use of raw MT.

To the best of my knowledge, no previous research has focused on how an environment in which negotiation for meaning is a common practice influences MT gisting. Considering that this dissertation contains the first in-depth examination of the use of MT by an ecosystem, this makes sense. The factor is firmly embedded in the environment in which MT use takes place, and capturing it requires qualitative study of specific environments.

6.3.4 Raw MT was used in a risk-tolerant environment

The final environmental factor that was found to influence the use and reception of raw MT was the environment's tolerance for risk. Article II analyzed the various risks in patent professionals work and how the IPR environment's tolerance for risk enhanced the possibilities to rely on raw MT for information. For a thorough discussion on this, see Section 5.2.1.2.

In contrast, Article V examined the idea of using MT for humanitarian purposes, but emphasized that an integral element of the idea involves "mitigat[ing] risks and negative impacts", as established in the principles of the IEEE Global Initiative for Ethical Consideration of Artificial Intelligence and Autonomous Systems (IEEE

2017, 23). The article examined processes that tend not to be risk-tolerant and therefore are not appropriate for raw MT use, namely, health and medicine. In a discussion on the risk of misinformation when users trust fluent yet incorrect translations, the authors concluded that, "Particularly in situations where a misunderstanding could lead to harm for the person accessing the information, or for others, raw MT alone is not sufficient" (Nurminen and Koponen 2020, 161).

6.3.4.1 Related work

Several studies have discussed the role of risk or error tolerance (or non-tolerance) in processes where MT is used. Way (2013, 6) cautioned about the use of raw MT for translating stock reports due to "potentially huge losses that could arise from mistranslation." Nitzke, Hansen-Schirra and Canfora (2019, 242) encouraged the consideration of potential "negative consequences" when deciding on whether to use MT or not, specifically pointing to consequences in the processes in which translated information is used. A few recent articles have given overviews on the use of MT in legal (Scott and O'Shea 2021) or legal and medical settings (Vieira, O'Hagan and O'Sullivan 2020) and highlighted the need for caution when using MT in critical processes. Martindale (2020) outlined three components that contributed to rendering the use of raw MT dangerous. These included erroneous yet believable MT output, a lack of means or motivation to verify the correctness of MT output, and the use of the machine-translated information to initiate actions or decisions.

On the other hand, researchers have also noted that certain contexts may be more tolerant of risk and that risk can be managed. Way (2018, 170) claimed that "Of course, some objectives could be more tolerant of MT errors than others," although unfortunately, he did not give examples of such objectives. Proposals have been made for applying risk management to MT scenarios (for example, Canfora and Ottmann 2020, Martindale 2020). And, as O'Brien (2020, 313) points out, in certain situations such as crisis communication, the risk of using MT needs to be balanced against the choice of giving no information at all: "Perhaps the most salient ethical question pertaining to translation and disaster settings is: what are the consequences of no translation?"

A few empirical studies have examined the topic of risk tolerance in processes. Yamashita and Ishida (2006a) reported that the use of raw MT disrupted conversants' indications that they understood their fellow conversant's messaging. This in turn led to them waiting for further input, thus adding to the time needed to satisfactorily conduct conversations. Yamashita and Ishida suggested that this

tendency made the process unsuitable for conversations that are very important, such as negotiations. Pituxcoosuvarn et al. (2018) studied ways to make communication processes more risk-tolerant, investigating the methods children involved in MT-mediated conversations turn to when MT communication fails.

6.4 Contextual influences in conclusion

This chapter presented 11 contextual factors that were found to influence the use and reception of raw MT in this dissertation, referred to as contextual influences. Many of the factors have been studied or discussed in previous research as well, and all relevant studies were included in the discussion on each factor. In the next chapter, the contextual influences are compiled into a framework and proposals are made for possible applications of the framework.

7 RESULTS, RESEARCH QUESTION 3: THEORETICAL FRAMEWORKS FOR CONCEPTUALIZING THE PHENOMENON OF MT GISTING

My third objective for this dissertation involved exploring possible theoretical frameworks through which MT gisting could be conceptualized. I began the work with one preliminary hypothesis concerning the phenomenon: that certain factors in the contexts in which people use raw MT influence its use and reception. During the dissertation project, two other possible ways of conceptualizing raw MT use emerged, namely, through the theoretical lenses of risk management and distributed cognition.

These three approaches offer frameworks through which MT gisting can be viewed and analyzed. Depending on context and goals, one framework might prove more useful and appropriate in a specific situation than the others. For example, a study that compares the reception of human translation with machine translation would benefit from using the same framework for analyzing each. Also, since the study of MT gisting is in its early stages, alternative ways to conceptualize the phenomenon can accelerate future research by providing different options for approaching questions related to MT gisting.

The frameworks are not necessarily mutually exclusive. They overlap to some degree, and multiple frameworks could be applied to examine the same phenomenon. For example, when a user of raw MT accesses multimodal aspects of the source document to complement their understanding of the MT output of the source text, that can be seen as the user benefitting from contextual influences in their context of use, as the user mitigating risk by attaining the best possible understanding of a text, or as the user interacting with elements of their distributed network to enhance their cognition of the MT output.

While these frameworks will be valuable in forwarding study on the users of raw MT, they are useful beyond the research field. Currently numerous companies, governmental and non-governmental organizations, language service providers, and individuals are engaging with raw MT, and many more are hoping to expand their

linguistic reach through MT gisting in the future. Many of these parties have an interest in evaluating possible uses for raw MT or building processes to accommodate it, and they need information and tools to help them carry out such activities. The frameworks described here can help. In the next three sections, I describe each approach. In the final section I explore various processes in which the frameworks might be applied.

7.1 MT gisting and contextual influences

Chapter 6 introduced 11 factors in the context of use of MT gisting that were found to influence the use and reception of raw MT, which I categorized according to the categories included in the definition of *context of* use in ISO standard 9241-210:2019 (ISO 2019), namely, user qualities, user goals and tasks, and environment. This categorization could also be viewed as a framework through which MT gisting can be conceptualized and analyzed. Such a framework would be useful in the study of MT gisting for three reasons. First, although individual contextual influences have been identified in various past studies, no attempt has been made to compile them into one framework. As a result, their collective importance has not been appreciated and higher-level discussions on the role of context in raw MT reception is lacking.

A second benefit of the framework is that it brings the findings of research in the twin fields of MT for assimilation and MT for communication into the same framework and discussion. Although these branches of research share many characteristics, as described in Section 2.3, they have grown in separate spheres and have not achieved the mutual benefit that might be possible if we viewed them as different manifestations of the same phenomenon. Bringing them together into a single framework will also contribute to raising awareness of the role and importance of context in all raw MT reception.

A third point of importance is that the framework can be applied in future research on the influence of context in MT gisting. The list of contextual influences presented here is not meant to be exhaustive, but rather the start of a more comprehensive list that will be constructed over time as we build our knowledge of MT gisting. It is possible that, as research in the area grows and new factors are found, the framework may change or be discarded. However, if the discussion advances our understanding of the role of context in raw MT use, the framework has served its purpose.

I propose that the framework of contextual influences can provide a way to analyze contexts, either potential or existing contexts, of MT gisting. Each factor in the model is not meant to apply to every case of MT gisting; in any given context, one, a few, or even many of the elements may be present. The presence of multiple influences might indicate an increased possibility for successful use of raw MT. For example, if a user has good knowledge of the source language, uses a tool that shows them both the source and target texts side-by-side, and also is working with a colleague with whom they can collaborate on understanding the raw output, that user would have a heightened chance of using the raw MT effectively. Conversely, if there is only one factor present but it is strong, that might compensate for the lack of other contexual influences. For example, the patent professionals described in Articles II and III often did not have any knowledge of the source languages of the machine-translated texts they read. However, their deep understanding of patent document genres and the subject matters they worked with made it possible for them to successfully use raw MT. Also, certain factors might have a tendency to be present in specific types of use cases. For example, more environmental factors may be present in cases in which MT is used in an ecosystem, since the ecosystem may have built allowances for raw MT use.

7.2 MT gisting as risk management

Borrowing Pym and Matsushita's (2018, 1) statement that "[t]he decisions that translators make while translating can be studied as a particular form of risk management," I propose that the same approach can be applied to MT gisting. I base this position on the results described in Article II, which found that patent professionals often employ an assessment of risk, reflective of risk management princples, when making the important decision of whether to rely on their understanding of the raw machine translation of a patent document or to order human translation. For more details, see Section 5.2.2.

In fact, the process described by the patent professionals in the study can be mapped onto the description of risk management in ISO standard 31000:2018 and specifically, the core processes of defining the scope, context and criteria; risk assessment; and risk treatment (ISO 2018, 14-18). Informants of the study displayed an understanding of the parameters of the risks they undertook when relying on raw MT. First, they displayed an understanding of the scope, context and criteria at play in decisions on relying on raw MT. Next, they identified and evaluated risks,

weighing the costs and time needed for human translation against their trust in their own understanding of raw MT. The main risk was relying on raw MT that contained mistakes, resulting in consequences of loss of money, time, and reputation. Finally, patent professionals discussed different treatments used to mitigate risks in their environment. One example given was consulting with various parties, both internal and external, on any passages in the raw MT that were not fully understood.

Risk management therefore provides a second framework for analyzing and conceptualizing the phenomenon of MT gisting in the IPR context. We can assume that the patent professionals in this context employed more sophisticated risk assessment processes than those that other raw MT users rely on, as patent professionals constitute a professional group that has collectively relied on MT gisting for a considerable period of time. However, I propose that the framework of risk management may be useful for examining and understanding the behavior of other users of raw MT as well.

Conceptualizing MT gisting through this framework provides us with a way of analyzing people's actions as they use raw MT, which can offer a new understanding of the phenomenon. Pym's description of the benefit of risk management in studying translator processes would seem to fit MT gisting as well: "The categories of risk management invite studies that assess the strategies of all participants in a translation event and use that matrix to try to explain translator decisions" (Pym 2021, 453).

Another benefit could be gained through educating users to view MT gisting as risk management. First, it would encourage them to see it as an activity that can and should be managed through a process, whether that is a more informal personal thought process, or a formal risk management procedure used throughout an organization. It would promote the awareness that risks can be identified and managed. When people view risk as a vague and unknown entity, they may be deterred from using raw MT altogether. Encouraging them to identify specific risks and devise strategies to manage them would promote users' own agency in MT gisting. Second, approaching MT gisting through the lens of risk management would bring a more detailed structure to users' analysis of the various risks involved in using raw MT and would encourage them to also consider the benefits to be gained. In organizational use, risk management could help to identify processes that could be developed to better withstand risk.

7.3 MT gisting as distributed cognition

The conclusions of Article III were that distributed cognition was an appropriate framework for studying the use of raw MT by patent professionals in the IPR ecosystem, and that viewing the phenomenon through this framework helped to build on our understanding of raw MT reception through a nuanced account of MT gisting in the ecosystem. In fact, the idea that "[u]nderstanding is an activity that crucially depends on the environment – and also on experience – because environmental affordances foster and constrain meaning construal" (Muñoz Martín 2017, 564) fully supports one of the conclusions of this dissertation, namely, that the context of use of raw MT influences the ways users engage in MT gisting and their reception of raw MT. It is therefore logical that distributed cognition was a useful approach to analyzing MT gisting.

The first benefit of analyzing MT gisting through the framework of distributed cognition is that it brings a deeper level of understanding to the phenomenon. When applied to the case of patent professionals' use of raw MT, it helped in identifying the various actors, both human and non-human, that contributed to patent professionals' cognitive efforts. It effectively highlighted how interaction with these elements facilitated the task of understanding raw MT and contributed to making the use of raw MT tenable.

Another benefit of applying distributed cognition to MT gisting is that it contributes to the small body of existing research on the cognitive aspects of MT gisting (see Section 3.6). Distributed cognition, as well as other situated approaches, open new pathways for studying cognition in MT gisting. Since these approaches most often involve studying MT users 'in the wild,' this might also catalyze researchers to augment laboratory-based studies with new methods.

A third benefit of applying the framework of distributed cognition to MT gisting is that it can contribute to theory-building in cognitively oriented Translation Studies more generally. As explained in Article III, much of the theoretical background I relied on when studying MT gisting through the lens of distributed cognition came from theoretical work on situated cognition and the translation process. This was necessary because there was a lack of theoretical thinking in relation to the reception of raw MT. It is possible that this approach to the study of MT reception might also be useful for researchers studying cognition in the reception of human translation or in the translation process itself.

Other situated approaches to cognition (for example, embodied, embedded, or extended) might also be applied to the study of users of raw MT. Distributed

cognition was a logical framework in the study on patent professionals because they interact with such a wide network to achieve understanding of raw MT. In describing how distributed cognition differs from other approaches, Risku and Rogl (2021, 487) stated that, "the focus in distributed cognition lies more on the distributedness of cognition and action." However, other approaches might also be fitting for this use case or other cases of MT gisting.

7.4 Application of the frameworks

This chapter has suggested three frameworks through which the phenomenon of MT gisting might be viewed and studied, including a framework of contextual influences, risk management, and distributed cognition. All three of these frameworks can be useful in defining and developing future research on MT gisting. However, they could also be applicable and useful in more practical areas in which raw MT is currently being used or is under consideration. The following sections introduce five different areas where these frameworks could be applied. Examples of potential use cases are presented for each.

7.4.1 The frameworks in future research

Analyzing the phenomenon of MT gisting through the frameworks of contextual influences, risk management, or distributed cognition can help researchers understand raw MT reception in a more nuanced way and to develop a deeper understanding of all factors that contribute to the context of use. It can also help to identify open questions and areas for further research of different types.

The framework of contextual influences provides a mapping of a number of paths for further experimental research. Many of the factors identified in this thesis had been identified and empirically tested in previous research. However, more comprehensive testing of each one would be warranted. The relative importance of each factor in relation to the others could also be tested. For example, many studies have compared the output of different methods of translation: human, post-edited, or raw MT (for details, see Section 3.2.2). Similar comparisons could be conducted that would test, for example, if source language knowledge affects MT gisting more than familiarity with context. Finally, the current list of contextual factors is in no way meant to be complete and static. Future research could also focus on identifying

and studying new contextual factors that influence MT gisting. For example, users' attitudes towards MT or their tolerance for uncertainty might prove to be important influences on their reception of raw MT.

Further research on patent professionals could assess how they view and employ risk assessment in their use of raw MT. Surveys could be used to access a wider variety of informants, including those from different geographical areas, with different linguistic backgrounds, or with different levels of experience in IPR processes. Another area to be investigated is how risk is accounted for by other groups of people as they use raw MT. For example, it would be interesting to obtain a detailed account of how users of free online MT view the risks involved in using the tools.

Finally, this thesis examined how patent professionals' use of MT could be viewed as distributed cognition. Future studies could investigate this concept further in the context of patenting work. It would also be interesting to investigate other professional contexts or ecosystems in which MT gisting takes place in order to evaluate whether the use of distributed cognition is as prevalent. Also, in the patent professional as well as other use cases, alternative approaches to situated cognition might also be applicable, including embedded, embodied, extended and affective cognition. I return to the topic of future research later in Section 8.3.

7.4.2 Using the frameworks to evaluate the suitability of raw MT for different use cases

Researchers, organizations, individuals, and language service providers that also provide consulting on language issues all have an interest in evaluating whether specific use cases – documents, text types or contexts – might be suitable places to use raw MT. Past research has proposed the characteristics that might be present in good cases for MT, including aspects of the proposed users, the texts to be translated, and abstract factors such as purpose and value (see Section 3.4). The inclusion of more specific aspects of context in evaluation processes could result in more comprehensive evaluations and more effective predictions of good (or poor) contexts for raw MT use, especially if *context* does not remain a vague concept but rather is described through well-defined contextual factors such as those introduced in Chapter 6.

Risk management could also provide a comprehensive lens for evaluating potential use cases for raw MT. Technical, business and content-related evaluation

factors all contain risks. Technical risks might involve underdeveloped tools or language pairs that are not well supported yet. Business risks include both risks to people, animals and equipment and risks to the business's reputation. The content to be translated might also be considered risky, for example, if it was not authored in a standardized way and therefore could easily be mistranslated. Risks in the context of use would also apply. For example, risks in the processes in which people use raw MT should also be considered. A robust evaluation based on risk management principles would acknowledge and include the risks from all of these areas instead of only one or two. It would also provide a framework for assessing those risks, deciding on treatments, and making decisions.

To give an example of a practical application in the research world, I offer the situation I faced at the beginning of the study that led to Article IV. I had access to important informants whose input would greatly increase the scope and impact of my research. However, the informants were in a different geographical region and they and I did not share a language we could use in interviews. I did not have access to information on how to evaluate whether I could successfully conduct the interviews using MT-mediated communication, so I chose to simply try it.

A researcher who is contemplating such a choice in the future, however, could use the framework of contextual influences to evaluate this choice. They might first examine how much the interviewer and informant know of each other's languages. They could then consider how deeply the informants know the subject matter to be discussed and how familiar MT is to them. After contemplating all factors, they might consider whether anything could be done to enhance the help to be gained from contextual factors. In this example, they might identify two things. The initial interview questions could be professionally translated and shared with informants ahead of time so that the informants would be better acquainted with the subject matter, and the informants' level of MT literacy could be enhanced before the interview by sending them information on MT. Armed with all of this information, the researcher could then decide on how to proceed.

7.4.3 Using the frameworks in product development

MT as a technology is starting to gain maturity. It is already good enough for millions of people to use for numerous purposes. For developers of MT tools, this means that offering MT is not enough. To differentiate themselves from the competition, they need to think beyond the MT engine and consider adding other features that

help end users and enhance their experience with the technology. This issue was noted already in the later stages of the development of the last generation of MT technology, statistical MT. In a study, Gaspari (2007, 240–241) chose to "concentrate on issues of usability and HCI [human-computer interaction] that we believe are much more likely to lead to noticeable improvements of benefit to the users of online MT, especially in the short term." The findings of this dissertation similarly indicate that there are many ways to enhance the experience of MT gisting. Developers of MT tools could use the framework of contextual influences or risk management to identify auxiliary features and tools that could enhance people's use of raw MT. Following is a fictitious example of how this might be applied.

Company Y has a product which aggregates different MT engines into one user interface and provides a way to embed that user interface into a company's own intranet, offering the company's information to global employees in a variety of languages. Company Y does not develop any of the MT engines used, they only build the user interface environment that pulls them together. Company Y's software team is interested in developing additional features and tools that would enhance the user experience and give them a competitive advantage over companies producing similar solutions. Company Y reviews the framework of contextual influences to better understand the elements that have been found to enhance users' experience with MT. They identify the following potential tools and features: a user interface that easily shows both source and text side-by-side; a feature that shows users the output from two or three of the MT engines, allowing them to select the best one; and messages that appear in strategic locations of the user interface to push information and tips on how to use MT effectively.

The software team might also approach the same issue from the viewpoint of risk management, aiming to implement features that reduce risk or encourage users to identify and manage risk. For example, they could introduce an automated quality estimation feature that would mark pieces of text that score poorly, thus helping users to identify potentially risky information and to take extra care in verifying the translation, as suggested by Martindale (2020, 42), or in interpreting its meaning. Implementing a feature that reveals automated quality estimations to users of raw MT has been suggested by Martindale (ibid.), Specia and Shah (2018, 202) and Lavie (in Campbell et al. 2020).

7.4.4 Using the frameworks in process development

Organizations that use or disseminate raw MT texts, as well as the language service providers that advise them, can use the frameworks of contextual influences and risk management to develop processes that accommodate MT gisting. Assessing the contextual factors that affect raw MT use in their own environments could be a starting point for planning how to develop their processes or introduce new technology. Analyzing their own organization and processes through the lens of risk management could help them to first pinpoint process phases that carry higher risk, then to find strategies for avoiding, sharing, or reducing that risk.

The following fictitious example from the business world was inspired by a question posed in a social media forum on machine translation. Company X is a software company with engineering teams in Japan and the U.K. Most of the Japanese engineers have some proficiency in English, though not enough to use it in their work. Currently much of the documentation they need to ensure they work consistently with the teams in the U.K., for example product specifications and code comments, is not translated into Japanese. Rather, a few team members who have a better command of English read these documents and explain them to the others.

Company X has decided to start providing the documents to the Japanese engineering team as raw MT and employs a risk management framework to help them identify potential risks and develop their processes to be more risk tolerant. One of the risks they identify is that a mistranslation in the machine-translated documentation could lead to one team producing a piece of software that is not compatible with other parts of the software product made by the other team. They further analyze this risk, evaluating potential consequences and the likelihood of the risk occurring. They then weigh these considerations against their organization's own principles of what level of risk they are willing to take. Finally, they decide on treatment. In this example, they decide to reduce the risk by implementing more frequent cross-site software reviews, which would quickly identify erroneous actions and prevent large losses of time.

7.4.5 Using the frameworks to develop MT literacy

The frameworks introduced in this section can also contribute to increasing MT literacy, building MT literacy training programs, and producing guidelines for users of raw MT. Specifically, the frameworks give valuable input for building literacy around the third and fourth components of the definition proposed in this dissertation (see Section 2.4.1), which concern an MT user's ability to:

3. Understand how machine translation systems are or can be used for purposes that are important to the user
4. Appreciate the wider implications associated with the use of MT

A better understanding of the role of context in raw MT reception contributes to our comprehension of how MT is being, and can be, used for purposes that are important to users. A better understanding of the contextual factors that influence MT use and reception can help people identify potential use cases for MT gisting and evaluate whether MT might, or might not, be suitable for those purposes. Risk management helps us understand how users view the risk involved in different use cases and gives us insight into their behavior. Finally, viewing MT gisting through the lens of distributed cognition can help us understand some of the current use cases for MT and identify potential new use cases, which might be overlooked if the situated nature of MT use is ignored. The fostering of this type of understanding would fit naturally into MT literacy training programs. It would help students learn to evaluate possible uses of raw MT in specific contexts, making decisions on when and where the use of raw MT might be appropriate and where post-editing or human translation would give a better result.

A training module on the fourth component of MT literacy would benefit from an approach based on risk management. The training would encourage people to view the use of MT as something that should be approached through risk management. It would give them tools and processes that help them identify risks and possible consequences, evaluate benefits, identify possible treatments, and make decisions. It would encourage students to work with risk where tenable, but to use other solutions than raw MT in cases where the risk is too great.

Finally, besides formal training, MT literacy can be increased through guidelines for users. Effective guidelines would need to contain useful information and be presented to users in appropriate places and at appropriate times. Producers of such guidelines can use the contextual factor framework to outline behaviors that they

might want to encourage users to adopt. For example, guidelines could encourage users to also refer to the source text if they have any proficiency in source language or to check unclear passages by putting them text into a second MT tool. The risk management framework can provide input on the type of information users should be given concerning the risks involved in using raw MT.

The frameworks described in this chapter, and the examples of practical applications of them, provide structures and conceptualizations for further academic research, but could also be used by industry. It is hoped that the insights on contexts and end users of raw MT produced by this dissertation work will inspire researchers and developers of MT tools to look beyond the MT technology to find new ways to enhance the experience of people using MT for gisting.

8 CONCLUSIONS

The first goal of this dissertation was to examine a limited number of specific contexts of MT gisting. A total of four contexts were analyzed, including online MT use, MT gisting within a professional ecosystem, interviewing using MT-mediated communication to gather data for research, and situations in which people can benefit from better accessibility to the information they need to participate fully in society. In the area of online MT, a study was conducted of the users of one online MT tool. Data from logs and a survey revealed that the tool was used by a geographically dispersed and diverse group of users. Respondents reported using MT most often for gisting purposes, for their own understanding, someone else's understanding, or to verify that they've understood a text. The most common area of life respondents used raw MT in was study, followed by work and leisure. Finally, when asked how well they understood the language of the document they translated using MT, 83% responded that they understood the language a little, well, or very well. The study concluded that this possibly meant that they view the use of MT as a somewhat different thing than the way they use human translations.

The second context examined was one in which a group of professionals used raw MT in a business ecosystem and according to the conventions and practices afforded MT gisting by that ecosystem. Patent professionals employed MT gisting to access patent documents which they needed in their work, but which were in languages the patent professionals did not speak. The use of raw MT in the environment was transparent, considered legitimate, and supported by guidelines. The context was described as one with many inherent risks, and the risky nature of the field contributed to making MT gisting tenable. First, it led to a tolerance for risks of all types, including the risks involved in using raw MT. Second, the patent professional informants displayed a tendency to employ risk assessment when using MT and when deciding between relying on raw MT or sending documents for human translation. A separate analysis of the study data concluded that patent professionals' practice of MT gisting could also be analyzed and understood through the lens of distributed cognition, and that such an analysis produced a detailed and nuanced account of the practice.

In the third context investigated in the dissertation, an academic researcher interviewed informants with whom she did not share a languge. Communication took place in an MT-backed chat application, with interviewer and informants each typing messages in their own languages and those messages being machine-translated. This method of interviewing via MT-mediated communication was deemed an interesting and promising method for enabling the inclusion of more diverse and geographically dispersed participants in acadmic research.

Finally, the dissertation examined research and projects that concern the contexts in which people suffer from difficulties in accessing information in the societies they live in, including people who have recently moved there, refugees, and asylum seekers. A number of efforts to implement MT to increase accessibility and enable better participation in civil life and better access to health, culture and media have been studied or piloted. While many results are promising, there are also challenges to be overcome, including ethical questions concerning high-risk content, varying levels of quality in MT, and the need to involve targeted groups in development.

The dissertation's second goal was to identify factors in the studied contexts that were observed or reported to influence the use and reception of raw MT. A total of eleven factors were identified and these were categorized as factors relating to users, to their tasks and goals, or to factors found in the technial and organizational environments. The factors that related to the users included their language competences (source, target or pivot languages), how familiar they were with the textual context of the material being machine-translated, and how high a level of MT literacy they possessed. Factors related to user tasks and goals involved actions and processes they employed while using MT to maximize their understanding of the raw MT. Users accessed pictures and other multimodal elements in the source texts, they put the same material into two different MT tools, they adapted messages to get improved MT output, and they negotiated the meaning of raw machine-translated texts with others. Environmental factors that affected raw MT use and reception included auxiliary technologies that were used alongside MT, the status and acceptability of the practice of MT gisting, and supports for using raw MT by environments in which negotiating the meaning of texts is common and accepted, and risk is tolerated.

Finally, the dissertation aimed to explore ways to conceptualize the phenomenon of MT gisting. Three different frameworks that could be used to understand and analyze MT gisting were proferred, including the framework of contextual factors that influence raw MT use and reception, the framework of risk management and the framework of situated cognition. A number of suggestions were offered for

applying the frameworks in academic and industrial settings. For example, the frameworks could be useful in future research, when evaluating the suitability of raw MT in various use cases, in process or product development, or when developing MT literacy programs.

8.1 Contributions

8.1.1 Insights into some contexts in which raw MT is used

The dissertaton examined a diverse group of contexts in which MT gisting takes place, including MT use in open contexts such as the internet, closed environments such as the patenting ecosystem and academic research, and public service contexts in which information is produced for broad distribution. This diversity led to a wide array of insights, many of which are specific to the individual contexts. One common finding was that contextual factors were identified that were found to affect the way people used and received raw MT; this common finding is discussed in Section 8.1.2.

Article I offered the first survey study of online MT use since those conducted by Gaspari (2007) and Burgett (2015), and therefore brought an updated view to online MT use. It was the first to occur after the introduction of neural MT, although NMT was most likely only available for some of the language pairs included. The study highlighted some previously unreported trends in MT gisting, such as the popularity of MT gisting for study and a tendency to translate texts that are in languages people already have some understanding of.

To the best of my knowledge, this dissertation's research on MT gisting in the IPR ecosystem was the second qualitative study ever conducted on the use of MT by non-translator professionals, trailing the first study (Henisz-Dostert 1979) by 40 years. The first contribution of the study was the attention it brought to the widespread use of raw MT in this professional environment. Although MT gisting had been a common practice for ten years (Nurminen 2019, 32) at the time of the study, it had been unrecognized by the Translation Studies research community. In fact, there was little recognition of MT gisting by any professional group other than translators, nor did we have an awareness that MT gisting can be legitimate and professionally conducted by a group. The study's second contribution was a deep description of MT gisting in the ecosystem. Such concentrated use of raw MT provided an effective space for observing phenomena that might also apply to MT

gisting on a more individual level. Third, the study provided insights to the shared ideas of legitimacy, guidelines (both formal and informal), and training that contribute to making the use of raw MT tenable in such an environment.

The pilot investigation into gathering data for research by interviewing participants using MT-mediated communication provided insights into the possibilities the method offers, while the examination of the possible use of MT for increasing accessibility to information contributed insights to the ongoing conversation on ethics and MT. It highlighted the idea that MT can provide benefits to humanity and that the quest to apply technologies to the improvement of lives is an integral part of ethics.

8.1.2 Context affects use and reception

The dissertation provided evidence that certain aspects of the context of use in which MT gisting occurs influence how people use raw MT and how they receive it. In past research, the effect of individual aspects of context had been studied, but a compiled view and discussion of context's effect on reception was missing. The framework of eleven contextual factors in this dissertation serves to both highlight the influence context has on reception and to define specific contextual factors that been found to have an effect. Past studies that have defined characteristics of good use cases for MT (see Section 3.4) have considered characteristics related to technology, business requirements, the content to be translated, and some aspects of context. This dissertation contributed new insights into contextual features that can contribute to making a use case for MT gisting successful or not. Finally, the combining of contextual factors from past studies on MT for assimilation and MT for communication into one framework provides a previously missing viewpoint that they are two different manifestations of the same phenomenon of MT gisting.

8.1.3 Conceptual and terminological contributions

Although the first study of MT gisting completed in 1979 (Henisz-Dostert 1979), the scarcity of research over the years has led to a lack of conceptual development. This dissertation contributed a discussion on both the conceptualization of the overall phenomenon of MT gisting and on some of the individual terms and concepts we use when discussing it. The suggestions that MT gisting could be conceptualized as an activity which is supported by contextual factors, or as an

exercise of risk management or distributed cognition, contribute three ways of viewing, studying and discussing MT gisting. To the best of my knowledge, this may be the first of any type of theoretical framing of MT gisting beyond a definition of terms or description of use. In addition to this, several developments to existing terms and concepts were suggested in the dissertation, including the first comprehensive definition of *MT gisting* that is not meant to be applied to only one individual use case*;* a revisitation and proposed revision of the categories of use that was proposed by Hutchins more than 30 years ago (Hutchins 2010, 29) and has been used without changes since; and two additions to Bowker and Buitrago's (2019) definition of the concept of *MT literacy*.

8.1.4 Methodological contribution: MT-mediated interviewing

In one study of this dissertation, I piloted the use of MT-mediated communication for gathering data for research. Interviews were conducted between an interviewer and informants who did not share a language, and the medium of communication was a chat application with integrated MT. The method was found to be tenable enough to warrant further research, and several considerations were offered for researchers who might consider using the method. Its benefits include the opportunity to reach a more diverse group of potential study informants without additional costs such as travel and interpreting services.

8.2 Limitations of the research

The first limitation in the studies that made up this dissertation is that several of them involved limited data sets. The study on patent professionals involved nine informants who were similar in geography, level of experience, and age. The study that led to Article IV involved 4 interviewees and myself. However, these studies were qualitative and exploratory. It is the nature of qualitative research to have a narrow, yet deep, focus. This desribed the study on patent professionals especially well. The study provided a deep look at MT gisting, through which various avenues of new inquiry could be identified. Nevertheless, it would be beneficial for future studies to include larger and more diverse groups of participants. Conversely, the study on users of an online MT tool included a large number of participants, but a limited scope of investigation, with log data that focused on countries of origin and

languages and a survey consisting of eight questions. A more comprehensive survey would have enabled a more well-rounded analysis of that particular user group and a deeper investigation of their online MT use.

Although the diversity of contexts examined in the work contributed a multi-faceted description of users and practices, a more focused look at one specific sector (for example, MT gisting in the workplace) might have produced a more robust analysis of that specific sector. Various methods could then be applied to examine different aspects of gisting, and a larger spectrum of the user population could be considered. The diversity of methods applied in the dissertation led to a similar limitation. Instead of applying sundry methods to a variety of contexts, employing the same method in the study of different contexts would have enabled a more robust comparison of those contexts.

Overall, it is important to note that most of the findings of this dissertation cannot be generalized. For the most part, they are based on informants' own reports of MT gisting by themselves and colleagues, and they must be understood as such. However, the findings provide groundwork for future studies which can empirically test and verify them.

8.3 Future directions

As stated in the introduction, this was meant to be a dissertation with a broad scope that would generate new questions and identify areas for future work. For this reason, although the following list of ideas for future work is relatively long, it represents only a few of the myriad paths researchers of MT gisting might follow.

This dissertation explored some unrecognized or under-studied contexts for raw MT use, including MT gisting in the business ecosystem of patent work, MT-mediated interviewing in research, and MT for increasing accessibility to information. Future studies could focus on existing or emerging contexts which have not been sufficiently studied yet. Examples of such contexts include customer support scenarios in which agents support customers with whom they do not share a language through MT-mediated communication, discovery processes in legal settings, academic work such as reading relevant literature as raw MT, and interviewing via MT-mediated communication.

It would be important to conduct further studies on the contexts examined in this dissertation. More comprehensive and granular surveys of online MT users would be warranted. The users of raw MT in the IPR world should also be examined

further, for example, by including a larger number of informants who represent different geographical regions and a wider variety of experience levels. It would also be interesting to further investigate how the IPR ecosystem achieved broad, transparent, and rule-guided use of raw MT. A better understanding could help stakeholders in emerging use cases, for example, in legal eDiscovery or pharmacovigilance ecosystems, to build the necessary training and processes to accommodate raw MT use. Finally, the further study of how MT can be used to increase accessibility to information would be an important line of inquiry.

The dissertation produced evidence of 11 contextual factors that influenced the use and reception of raw MT. Many of the factors had been identified in previous literature, and some had been tested in one or two empirical studies. However, further empirical evidence of these and other contextual factors, and their specific effect on reception, is needed. One group of studies in MT has focused on empirical testing of the effects of different translation types (human, raw MT or PE) on the reception of texts (see Section 3.2.2). Similar studies could compare the influence of different contextual factors on reception. For example, a study might address the question of which is more influential: competence in the source language or knowledge of the subject matter of a raw MT text? In addition to empirical studies of the identified contextual factors, research that identifies new factors would also be called for.

One proposal that resulted from this work is that MT gisting could be conceptualized as risk management. Future work could delve deeper into risk management and MT gisting, focusing on users' views of risk and risk management, risk management in MT literacy programs, or implementation of risk management practiced in different MT gisting use cases (see Koponen and Nurminen, forthcoming).

Another proposal was that MT gisting could be viewed as distributed cognition. Further studies could explore MT gisting in relation to other flavors of situated cognition, such as embodied, embedded, extended and affective cognition. Past studies (for example, Doherty and O'Brien 2014; Castilho et al. 2014) have used eye tracking to examine cognitive effort in raw MT reception. Similar methods could be used to establish effort in different contexts of MT gisting. Methods from cognitive studies might also be employed to examine how users fill in the gaps caused by unclear passages in raw MT, how effective they are at spotting errors, and what types of strategies they employ to overcome information that is unclear due to linguistic errors.

Finally, the study of the use and users of raw MT would benefit from the application of a wider variety of research methods. Surveys can continue to provide us with certain types of knowledge, but qualitative methods such as contextual inquiry or narrative inquiry might bring new insights to the phenomenon. Approaches such as ethnography or autoethnography could be applied to research on MT use in ecosystems.

I conclude this dissertation in the same way I have concluded many research articles. Namely, with the hope that the future will bring a growth in research on all aspects of MT gisting.

BIBLIOGRAPHY

Abdallah, Kristiina and Kaisa Koskinen. 2007. "Managing Trust: Translating and the Network Economy." *Meta* 52 (4): 673–687.

Ageeva, Ekaterina, Mikel L. Forcada, Francis M. Tyers, and Juan Antonio Pérez-Ortiz. 2015. "Evaluating Machine Translation for Assimilation Via a Gap-Filling Task." *Proceedings of the 18th Annual Conference of the European Association for Machine Translation*: 137–144.

Albrecht, Urs-Vito, Marianne Behrends, Regina Schmeer, K. Matthies Herbert, and Ute von Jan. 2013. "Usage of Multilingual Mobile Translation Applications in Clinical Settings." *JMIR mHealth uHealth* 1 (1).

Anazawa, Ryoko, Hirono Ishikawa, and Kiuchi Takahiro. 2013. "Use of Online Machine Translation for Nursing Literature: A Questionnaire-Based Survey." *The Open Nursing Journal* 7: 22–28.

Aymerich, Julia and Hermes Camelo. 2009. "The Machine Translation Maturity Model at PAHO." *Proceedings of the Twelfth Machine Translation Summit*: 403–409.

Bakardjieva, Maria. 2011. "The Internet in Everyday Life: Exploring the Tenets and Contributions of Diverse Approaches." In *The Handbook of Internet Studies*, edited by Mia Consalvo and Charles Ess, 59–82. Chichester, United Kingdom: Wiley-Blackwell.

Bar-Hillel, Yehoshua. 1964. "A Demonstration of the Nonfeasibility of Fully Automatic High Quality Translation." In *Language and Information*, 174–179. Reading, Mass.: Addison-Wesley.

Beninatto, Renato and Michael Stevens. 2019. *What's the Latest with Neural MT?* Globally Speaking Radio 75. RWS Moravia. Retrieved from: https://www.globallyspeakingradio.com/episodes/whats-the-latest-with-neural-mt/

Birch, Alexandra, Ondrej Bojar, Rudolph Rosa, Juliane Ried, Hayley Hassan, and Colin Davenport. 2017. *D5.5: Report on User Surveys, Impact Assessment and Automatic Semantic Metrics*: Health in my Language project.

Birch, Alexandra, Juliane Ried, Colin Davenport, Matthias Huck, and David Marecek. 2018. *D5.6: Report on Third Year's Evaluation*: Health in my Language project.

Blench, Michael. 2008. "Global Public Health Intelligence Network (GPHIN)." *Proceedings of the 8th Conference of the Association for Machine Translation in the Americas*: 299–303.

Bouillon, Pierette, Johanna Gerlach, Hervé Spechbach, Nikos Tsourakis, and Sonia Halimi. 2017. "BabelDr Vs Google Translate: A User Study at Geneva University Hospitals (HUG)." *Proceedings of the 20th Annual Conference of the European Association of Machine Translation*.

Bowker, Lynne. 2009. "Can Machine Translation Meet the Needs of Official Language Minority Communities in Canada? A Recipient Evaluation." *Linguistica Antverpiensia* 8: 123–155.

Bowker, Lynne. 2020. "Fit-for-Purpose Translation." In *The Routledge Handbook of Translation and Technology*, edited by Minako O'Hagan, 453–468. Oxon, UK: Routledge.

Bowker, Lynne and Jairo Buitrago Ciro. 2015. "Investigating the Usefulness of Machine Translation for Newcomers at the Public Library." *Translation and Interpreting Studies* 10 (2): 165–186.

Bowker, Lynne and Jairo Buitrago Ciro. 2018. "Localizing Websites using Machine Translation: Exploring Connections between User eXperience and Translatability."

In *The Human Factor in Machine Translation*, edited by Sin-wai Chan, 7–29. Abingdon, Oxon, UK: Routledge.

Bowker, Lynne and Jairo Buitrago Ciro. 2019. *Machine Translation and Global Research: Towards Improved Machine Translation Literacy in the Scholarly Community*. Bingley, UK: Emerald Publishing Limited.

Bowker, Lynne and Melissa Ehgoetz. 2007. "Exploring User Acceptance of Machine Translation Output: A Recipient Evaluation." In *Across Boundaries: International Perspectives on Translation Studies*, edited by Dorothy Kenny and Kyongjoo Ryou, 209–224. Newcastle, UK: Cambridge Scholars Publishing.

Brace, Colin, Muriel Vasconcellos, and L. Chris Miller. 1995. "MT Users and Usage: Europe and the Americas." *Proceedings of MT Summit V*.

Braun, Virginia and Victoria Clarke. 2006. "Using Thematic Analysis in Psychology." *Qualitative Research in Psychology* 3 (2): 77–101.

Braun, Virginia and Victoria Clarke. 2013. *Successful Qualitative Research: A Practical Guide for Beginners*. Los Angeles, Calif.: SAGE.

Braun, Virginia and Clarke, Victoria. "Thematic Analysis: A Reflexive Approach." last modified n.d., accessed March–June 2019. https://www.psych.auckland.ac.nz/en/about/our-research/research-groups/thematic-analysis.html.

Burgett, Will. 2015. "Unmoderated Remote Usability Testing of Machine Translation Content." *TAUS Review of Language Business and Technology* IV: 30–37.

Burgett, Will and Julie Chang. 2008. "The Triple Advantage Factor of Machine Translation: Cost, Time-to-Market and FAUT." *Proceedings of AMTA 2008*.

Burgett, Will, Chris Duran, Ray Fiorino, Valarie Gilbert and Martine Smets. 2012. "Many Tongues, One Community" Discussion Forum at TAUS User Conference 2012, Seattle.

Calefato, Fabio, Filippo Lanubile, and Rafael Prikladnicki. 2011. "A Controlled Experiment on the Effects of Machine Translation in Multilingual Requirements Meetings." *Proceedings of the Sixth IEEE International Conference on Global Software Engineering*: 94–102.

Calefato, Fabio, Filippo Lanubile, Tayana Conte, and Rafael Prikladnicki. 2012. "Assessing the Impact of Real-Time Machine Translation on Requirements Meetings: A Replicated Experiment." *Proceedings of the ACM-IEEE international symposium on Empirical software engineering and measurement*: 251–260.

Campbell, Janice, Alon Lavie, Milica Panic, Jaap van der Meer, Chris Wendt, and Andy Jones. 2020. *Automated Translation in the Enterprise Webinar*. Online: TAUS. https://www.taus.net/events/webinars/138-mt-series-automated-translation-in-the-enterprise

Canfora, Carmen and Angelika Ottmann. 2019. "Who's Afraid of Translation Risks?" In *Challenging Boundaries: New Approaches to Specialized Communication*, edited by Heike Jungst, Lisa Link, Klaus Schubert and Christiane Zehrer, 73–92. Berlin: Frank & Timme.

Canfora, Carmen and Angelika Ottmann. 2020. "Risks in Neural Machine Translation." *Translation Spaces* 9 (1): 58–77.

Carbonell, Jaime G., Alon Lavie, Lori Levin, and Alan W. Black. 2006. "Language Technologies for Humanitarian Aid." In *Technology for Humanitarian Action*, edited by Kevin M. Cahill, 111–138. New York, USA: Fordham University Press & The Center for International Health and Cooperation.

Castilho, Sheila. 2016. "Measuring Acceptability of Machine Translated Enterprise Content." Ph.D. dissertation, Dublin City University.

Castilho, Sheila and Sharon O'Brien. 2016. "Content Profiling and Translation Scenarios." *The Journal of Internationalization and Localization* 3 (1): 18–37.

Castilho, Sheila and Sharon O'Brien. 2017. "Acceptability of Machine-Translated Content: A Multi-Language Evaluation by Translators and End-Users." *Linguistica Antverpiensia* 16: 120–136.

Castilho, Sheila, Fabio Alves, Sharon O'Brien, and Morgan O'Brien. 2014. "Does Post-Editing Increase Usability? A Study with Brazilian Portuguese as Target Language." *Proceedings of the 17th Annual Conference of the European Association for Machine Translation*: 183–190.

Castilho, Sheila, Stephen Doherty, Federico Gaspari, and Joss Moorkens. 2018. "Translation Revision and Post-Editing: Industry Practices and Cognitive Processes: Industry Practices and Cognitive Processes." In *Translation Quality Assessment: From Principles to Practice*, edited by Joss Moorkens, Sheila Castilho, Federico Gaspari and Stephen Doherty, 9–38. Cham, Switzerland: Springer.

Castilho, Sheila, Joss Moorkens, Federico Gaspari, Iacer Calixto, John Tinsley, and Andy Way. 2017. "Is Neural Machine Translation the New State of the Art?" *The Prague Bulletin of Mathematical Linguistics* 108 (1): 109–120.

Church, Kenneth W. and Eduard H. Hovy. 1993. "Good Applications for Crummy Machine Translation." *Machine Translation* 8: 239–258.

Clark, Andy and David Chalmers. 1998. "The Extended Mind." *Analysis* 4 (1): 7-19.

Clark, Herbert H. 1996. *Using Language*. Cambridge, UK: Press Syndicate of the University of Cambridge.

Council of Europe. 1950. *European Convention on Human Rights (European Convention for the Protection of Human Rights and Fundamental Freedoms, as Amended by Protocols Nos. 11 and 14)*: Council of Europe.

Das, Prithwijit, Anna Kuznetsova, Meng'ou Zhu, and Ruth Milanaik. 2019. "Dangers of Machine Translation: The Need for Professionally Translated Anticipatory Guidance Resources for Limited English Proficiency Caregivers." *Clinical Pediatrics* 58 (2): 247–249.

DePalma, Donald A. 2007. "Limited English Proficiency Not a Bar to Citizen Access." *MultiLingual*, 46–50.

Dew, Kristin N., Anne M. Turner, Yong K. Choi, Alyssa Bosold, and Katrin Kirchhoff. 2018. "Development of Machine Translation Technology for Assisting Health Communication: A Systematic Review." *Journal of Biomedical Informatics* 85: 56–67.

Dillinger, Mike and Laurie Gerber. 2009. "Success with Machine Translation: Automating Knowledge-Base Translation, Part 1." *ClientSide News,* 10.

Doherty, Stephen. 2012. "Investigating the Effects of Controlled Language on the Reading and Comprehension of Machine Translated Texts: A Mixed-Methods Approach." Ph.D. dissertation, Dublin City University.

Doherty, Stephen and Sharon O'Brien. 2009. "Can MT Output be Evaluated through Eye Tracking?" *Proceedings of the Twelfth Machine Translation Summit*: 214–221.

Doherty, Stephen and Sharon O'Brien. 2012. "A User-Based Usability Assessment of Raw Machine Translated Technical Instructions." *Proceedings of the Tenth Conference of the Association for Machine Translation in the Americas*.

Doherty, Stephen and Sharon O'Brien. 2014. "Assessing the Usability of Raw Machine Translated Output: A User-Centered Study using Eye Tracking." *Intl. Journal of Human-Computer Interaction* 30: 40-51.

Durban, Chris and Alan Melby. 2008. *Translation: Buying a Non-Commodity*: American Translators Association. https://atanet.org/wp-content/uploads/2020/06/translation_buying_guide.pdf

Ehrensberger-Dow, Maureen. 2019. "Ergonomics and the Translation Process." *Slovo.Ru: Baltijskij Accent* 10 (1): 37-51.

EPO: European Patent Office. 2018. *Guidelines for Examination in the European Patent Office*. Munich, Germany: European Patent Office (EPO).

Ess, Charles and Mia Consalvo. 2011. "Introduction." In *The Handbook of Internet Studies*, edited by Mia Consalvo and Charles Ess, 1–8. Chichester, United Kingdom: Wiley-Blackwell.

European Commission Directorate-General for Communications Networks, Content and Technology (DG CNECT). 2020. *Report on the SME Consultation on eTranslation and Language Technologies*: European Commission.

Federici, Federico M. and Sharon O'Brien. 2020. *Translation in Cascading Crises*. London: Routledge.

Flanagan, Mary. 1996. "Two Years Online: Experiences, Challenges and Trends." *Proceedings of the 2nd Conference of the Association for Machine Translation in the Americas "Expanding MT Horizons"*: 192–197.

Flournoy, Raymond S. and Christopher Callison-Burch. 2000. "Reconciling User Expectations and Translation Technology to Create a Useful Real-World Application." *Translating and the Computer 22*: 1–8.

Flournoy, Raymond and Jeff Rueppel. 2010. "One Technology: Many Solutions." *Proceedings of AMTA 2010: the Ninth Conference of the Association for Machine Translation in the Americas*.

Forcada, Mikel. 2010. "Machine Translation Today." In *Handbook of Translation Studies, Vol. 1*, edited by Yves Gambier and Luc van Doorslaer, 215–223. Amsterdam: John Benjamins.

Forcada, Mikel, Carolina Scarton, Lucia Specia, Barry Haddow, and Alexandra Birch. 2018. "Exploring Gap Filling as a Cheaper Alternative to Reading Comprehension Questionnaires when Evaluating Machine Translation for Gisting." *Proceedings of the Third Conference on Machine Translation: Research Papers*: 192–203.

Fuji, M., N. Hatanaka, E. Ito, S. Kamei, H. Kumai, T. Sukehiro, T. Yoshimi, and H. Isahara. 2001. "Evaluation Method for Determining Groups of Users Who Find MT "Useful"." *Proceedings of the 8th Machine Translation Summit*.

Gao, Ge, Hao-Chuan Wang, Dan Cosley, and Susan R. Fussell. 2013. "Same Translation but Different Experience: The Effects of Highlighting on Machine-Translated Conversations." *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*: 449–458.

Gao, Ge, Bin Xu, Dan Cosley, and Susan R. Fussell. 2014. "How Beliefs about the Presence of Machine Translation Impact Multilingual Collaborations." *Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing*: 1549–1560.

Gao, Ge, Bin Xu, David Hau, Zheng Yao, Dan Cosley, and Susan R. Fussell. 2015. "Two is Better than One: Improving Multilingual Collaboration by Giving Two Machine

Translation Outputs." *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*: 852–863.

Gaspari, Federico. 2004. "Online MT Services and Real Users' Needs: An Empirical Usability Evaluation." *Machine Translation: From Real Users to Research 6th Confererence of the Association for Machine Translation in the Americas, AMTA 2014*: 74–85.

Gaspari, Federico. 2006. "The Added Value of Free Online MT Services: Confidence Boosters for Linguistically-Challenged Internet Users, a Case Study for the Language Pair Italian-English." *Proceedings of the 7th Conference of the Association for Machine Translation of the Americas*: 46–55.

Gaspari, Federico. 2007. "The Role of Online MT in Webpage Translation." Ph.D. dissertation, University of Manchester.

Gaspari, Fedrico and John Hutchins. 2007. "Online and Free! Ten Years of Online Machine Translation: Origins, Developments, Current use and Future Prospects (Article)." *Proceedings of MT Summit XI*: 199–206.

Gaspari, Federico, Hala Almaghout, and Stephen Doherty. 2015. "A Survey of Machine Translation Competences: Insights for Translation Technology Educators and Practitioners." *Perspectives* 23 (3): 333–358.

Girletti, Sabrina, Pierette Bouillon, Martina Bellodi, and Philipp Ursprung. 2019. "Preferences of End-Users for Raw and Post-Edited NMT in a Business Environment." *Translating and the Computer 41*: 47–59.

Guerberof, Ana, Joss Moorkens, and Sharon O'Brien. 2019. "What is the Impact of Raw MT on Japanese Users of Word: Preliminary Results of a Usability Study using Eye-Tracking." *Proceedings of MT Summit XVII, volume 1*: 67–77.

Hautasaari, Ari. 2010. "Machine Translation Effects on Group Interaction: An Intercultural Collaboration Experiment." *ICIC '10: Proceedings of the 3rd international conference on Intercultural collaboration*: 69–78.

Henisz-Dostert, Bozena. 1979. "Users' Evaluation of Machine Translation." In *Machine Translation*, edited by Werner Winter, 149–244. The Hague: Mouton Publishers.

Hoshino, Sadao. 1995. "Survey Report: Korya Eiwa: 100-Dollar Commodity has Expanded the Market." *MT News International: Newsletter of the International Association for Machine Translation* (12).

Hu, Ke, Sharon O'Brien, and Dorothy Kenny. 2020. "A Reception Study of Machine Translated Subtitles for MOOCs." *Perspectives* 28 (4): 521–538.

Hutchins, W. John. 1986. *Machine Translation: Past, Present, Future*. Chichester, England: Ellis Horwood Limited.

Hutchins, W. John. 2010. "Machine Translation: A Concise History." *Journal of Translation Studies* 13 (Special issue: the teaching of computer-aided translation, nos 1-2): 29–70.

Hutchins, W. John and Harold L. Somers. 1992. *An Introduction to Machine Translation*. London: Academic Press.

Iconic Translation Machines (Iconic). 2016. *MT Success Guide*. Iconic Translation Machines. Last accessed June 2021.

Iconic Translation Machines. 2020. *The Role of Neural Machine Translation in the Life Sciences*. Iconic Translation Machines. Last accessed June 2021. https://iconictranslation.com/industries/lifesciences/

IEEE: The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems. 2017. *Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems*: https://standards.ieee.org/content/dam/ieee-standards/standards/web/documents/other/ead_v2.pdf

Ishida, Toru. 2016. "Intercultural Collaboration and Support Systems: A Brief History." *PRIMA 2016: Principles and Practice of Multi-Agent Systems, 19th Annual Conference*: 3–19.

ISO (International Standardization Organization). 2018. *ISO 31000 Risk Management. Guidelines*. Helsinki: Suomen standardisoimisliitto.

ISO (International Standardization Organization). 2010. *ISO 9241-210:2010 Ergonomics of Human-System Interaction. Part 210, Human-Centred Design for Interactive Systems*. En Iso. Helsinki: Suomen standardisoimisliitto.

ISO (International Standardization Organization). 2019. *ISO 9241-210:2019 Ergonomics of Human-System Interaction. Part 210, Human-Centred Design for Interactive Systems*. En Iso. Helsinki: Suomen standardisoimisliitto.

Joho, Hideo, Leif Azzopardi, and Wim Vanderbauwhede. 2010. "A Survey of Patent Users: An Analysis of Tasks, Behavior, Search Functionality and System Requirements." *Proceedings of the Third Symposium on Information Interaction in Context: IIiX '10*: 13–24.

Jormakka, Noora. 2021. "Konekääntämisen Kompetenssit Kouluissa: Tarkastelussa Kieltenopettajien Käänninlukutaito" [Machine Translation Competences in Schools: Investigating Language Teachers' Machine Translation Literacy]. Master's thesis, Tampereen yliopisto.

Jönsson, Arne. 2016. *DigInclude - Digital Inclusion in the Networked Society for Groups with Special Needs (Proposal for Vinnova)*.

Kazmer, Michelle M. and Bo Xie. 2008. "Qualitative Interviewing in Internet Studies: Playing with the Media, Playing with the Method." *Information, Communicaiton & Society* 11 (2): 257–278.

Kenny, Dorothy, Joss Moorkens, and Félix do Carmo. 2020. "Fair MT: Towards Ethical, Sustainable Machine Translation." *Translation Spaces* 9 (1).

Khoong, Elaine C., Eric Steinbrook, Cortlyn Brown, and Alicia Fernandez. 2019. "Assessing the use of Google Translate for Spanish and Chinese Translations of Emergency Department Discharge Instructions." *JAMA Internal Medicine* 179 (4): 580–582.

Kirchhoff, Katrin, Anne M. Turner, Amittai Axelrod, and Francisco Saavedra. 2011. "Application of Statistical Machine Translation to Public Health Information: A Feasibility Study." *Journal of the American Medical Informatics Association* 18: 473–478.

Koehn, Philipp. 2020. Neural Machine Translation. Cambridge, UK: Cambridge University Press.

Koponen, Maarit and Leena Salmi. 2015. "On the Correctness of Machine Translation: A Machine Translation Post-Editing Task." *The Journal of Specialised Translation* January 2015 (23): 118-136.

Koponen, Maarit and Mary Nurminen. Forthcoming. "Risk Management for Content Delivery Via Raw Machine Translation." In *Communicating Across Languages in Times of Technological Change - Innovations in Translation and Interpreting Research, Practice and Training*, edited by Marion Winters, Sharon Deane-Cox, Svenja Wurm and Ursula Böser. Bloomsbury.

Koponen, Maarit, Brian Mossop, Isabelle S. Robert, and Giovanna Scocchera. 2021. *Translation Revision and Post-Editing: Industry Practices and Cognitive Processes*. Abingdon, Oxon, UK: Taylor and Francis Group.

Lawson, Veronica and Muriel Vasconcellos. 1993. "Forty Ways to Skin a Cat: Users Report on Machne Translation.". *Translating and the Computer 15*: 121–130.

List, Jane. 2012. "Review of Machine Translation in Patents - Implications for Search." *World Patent Information* 34 (3): 193–195.

Liu, Chao-Hong. 2018. "(LoResMT 2018) Technologies for MT of Low Resource Languages 2018." *Proceedings of the Workshop on Technologies of Low Resource Languages*.

Liu, Chao-Hong and Alina Karakanta. 2019. "(LoResMT 2019) The Second Workshop on Technologies for MT of Low Resource Languages." *Proceedings of The Second Workshop on Technologies for MT of Low Resource Languages*.

Liu, Nancy Xiuzhi and Matthew Watts. 2019. "Mobile Translation Experience: Current State and Future Directions." In *Impacts of Mobile Use and Experience on Contemporary Society*, edited by Xiaoge Xu, 193–212. Hershey, PA, USA: IGI Global.

Long, Michael H. 1983. "Native Speaker/Non-Native Speaker Conversation and the Negotiation of Comprehensible Input." *Applied Linguistics* 4 (2): 126–141.

Markham, Annette M. 2004. "The Internet as Research Context." In *Qualitative Research Practice*, edited by Clive Seale, Giampietro Gobo, Jaber Gubrium and David Silverman, 358–374. London: Sage.

Martindale, Marianna. 2020. "Responsible Gist MT use in the Age of Neural MT." *Proceedings of the 14th Conference of the Association of Machine Translation in the Americas. Workshop on the Impact of Machine Translation*: 18–45.

Martindale, Marianna J. and Marine Carpuat. 2018. "Fluency Over Adequacy: A Pilot Study in Measuring User Trust in Imperfect MT." *Proceedings of the 13th Conference of the Association for Machine Translation in the Americas (Volume 1: Research Papers)*: 13–25.

Matamala, Anna, Andreu Oliver, Aitor Álvarez, and Andoni Azpeitia. 2015. "The Reception of Intralingual and Interlingual Automatic Subtitling: An Exploratory Study within the HBB4ALL Project." *Translating and the Computer 37*: 12–17.

McDonald-Maier, Lisa. 2009. "Esp@cenet: Survey Reveals New Information about Users." *World Patent Information* 31 (2): 142–143.

Melero, Maite, Antoni Oliver, and Tony Badia. 2006. "Automatic Multilingual Subtitling in the eTITLE Project." *Translating and the Computer 28*: 1–18.

Merriam, Sharan B. and Elizabeth J. Tisdell. 2016. *Qualitative Research: A Guide to Design and Implementation*. The Jossey-Bass Higher and Adult Education Series. Fourth edition ed. San Francisco, CA: Jossey-Bass.

Miyata, Rei, Anthony Hartley, Paris Cécile, Midori Tatsumi, and Kyo Kageura. 2015. "Japanese Controlled Language Rules to Improve Machine Translatability of Municipal Documents." *Proceedings of MT Summit XV, vol. 1: MT Researchers' Track*: 90–103.

Miyazawa, Orie. 2002. "MT Training for Business People and Translators." *Proceedings of the 6th EAMT Workshop: Teaching Machine Translation*: 7–12.

Moberly, Tom. 2018a. "Doctors are Cautioned Against using Google Translate in Consultations." *BMJ* 363: Article K4546.

Moberly, Tom. 2018b. "Doctors Choose Google Translate to Communicate with Patients because of Easy Access." *BMJ* 363: Article K3974.

Monk, Andrew. 2009. "Common Ground in Electronically Mediated Conversation." In *Synthesis Lectures on Human-Centered Informatics #1,* edited by John M. Carroll, 1-45: Morgan & Claypool.

Morland, D. Verne. 2002. *Nutzlos, Bien Pratique, Or Muy Util? Business Users Speak Out on the Value of Pure Machine Translation*. Dayton, Ohio, YSA: NCR Corporation.

Muñoz Martin, Ricardo. 2010. "Leave no Stone Unturned: On the Development of Cognitive Translatology." *Translation and Interpreting Studies* 5 (2): 145–162.

Muñoz Martin, Ricardo. 2017. "Looking Toward the Future of Cognitive Translation Studies." In *The Handbook of Translation and Cognition*, edited by John W. Schwieter and Aline Ferreira, 555–572. London: Wiley-Blackwell.

Nielsen, Jakob. 1993. *Usability Engineering*. London, UK: Academic Press Limited.

Nitzke, Jean, Silvia Hansen-Schirra, and Carmen Canfora. 2019. "Risk Management and Post-Editing Competence." *The Journal of Specialised Translation* (31): 239–259.

Nomura, Saeko, Toru Ishida, and Mika Yasouka. 2003. "Open Source Software Development with Your Mother Language: Intercultural Collaboration Experiment 2002." *10th International Congress on Human-Computer Interaction*.

Nurminen, Mary. 2016. "Machine Translation-Mediated Interviewing in Qualitative Research: A Pilot Project." *New Horizons in Translation Research and Education* 4: 66–84.

Nurminen, Mary. 2018. "Machine Translation in Everyday Life: What Makes FAUT MT Workable?" TAUS eLearning Blogs. https://blog.taus.net/elearning/machine-translation-in-everyday-life-what-makes-faut-mt-workable

Nurminen, Mary. 2019. "Decision-Making, Risk, and Gist Machine Translation in the Work of Patent Professionals." *Proceedings of the 8th Workshop on Patent and Scientific Literature Translation*: 32–42.

Nurminen, Mary. 2020. "Raw Machine Translation use by Patent Professionals: A Case of Distributed Cognition." *Translation, Cognition & Behavior* 3(1): 100–121.

Nurminen, Mary and Niko Papula. 2018. "Gist MT Users: A Snapshot of the use and Users of One Online MT Tool." *Proceedings of the 21st Annual Conference of the European Association for Machine Translation*: 199–208.

Nurminen, Mary and Maarit Koponen. 2020. "Machine Translation and Fair Access to Information." *Translation Spaces* 9 (1): 150–169.

Nuutila, Pertti. 2005. "Rough Machine Translation in the Communication Process." Licentiate thesis, University of Tampere.

O'Brien, Sharon. 2010. "Controlled Language and Readability." In *Translation and Cognition*, edited by Gregory M. Shreve and Erik Angelone, 143–165. Amsterdam: John Benjamins Publishing.

O'Brien, Sharon. 2017. "Machine Translation and Cognition." In *The Handbook of Translation and Cognition*, edited by John W. Schwieter and Aline Ferreira, 313–331. London: Wiley-Blackwell.

O'Brien, Sharon and Johann Roturier. 2007. "How Portable are Controlled Language Rules? A Comparison of Two Empirical MT Studies." *Proceedings of MT Summit XI*: 345–352.

O'Brien, Sharon and Federico Marco Federici. 2019. "Crisis Translation: Considering Language Needs in Multilingual Disaster Settings." *Disaster Prevention and Management: An International Journal* 29 (2): 129-143.

O'Brien, Sharon and Maureen Ehrensberger-Dow. 2020. "MT Literacy–A Cognitive View." *Translation, Cognition and Behavior* 3 (2): 143–163.

O'Brien, Sharon, Rahzeb Choudhury, Jaap Van der Meer, and Nora Aranberri Monasterio. 2011. *Dynamic Quality Evaluation Framework*: TAUS Labs Report.

O'Brien, Sharon, Chao-Hong Liu, Andy Way, João Graca, André Martins, Helena Moniz, Ellie Kemp, and Rebecca Petras. 2017. "The INTERACT Project & Crisis MT." *Proceedings of MT Summit XVI:* 38–48.

Ogura, Kentaro, Yoshihiko Hayashi, Saeko Nomura, and Toru Ishida. 2004. "User Adaptation in MT-Mediated Communication." *The First International Joint Conference on Natural Language Processing (IJCNLP-04)*: 596–601.

Oladosu, John B. and Justice O. Emuoyibofarhe. 2012. "A Yoruba–English Language Translator for Doctor–Patient Mobile Chat Application." *International Journal of Computers and Applications* 34 (3): 149–156.

O'Mathúna, Dónal, Carla Escartín, Helena Moniz, Jay Marlowe, Matthew Hunt, Eric DeLuca, Federico Federici, and Sharon O'Brien. 2019. *Ethics Recommendations for Crisis Translation Settings*: INTERACT: The International Network in Crisis Translation.

Opdenakker, Raymond. 2006. "Advantages and Disadvantages of Four Interview Techniques in Qualitative Research." *Forum: Qualitative Social Research* 7 (4), Art 11.

Panić, Milica. "Implementation of Machine Translation: What to Consider?" Translation Automation User Society (TAUS), accessed 10 September, 2020. https://blog.taus.net/knowledgehub/implementation-of-machine-translation-what-to-consider.

Parra Escartín, Carla and Helena Moniz. 2020. "Ethical Considerations on the use of Machine Translation and Crowdsourcing in Cascading Crises." In *Translation in Cascading Crises*, edited by Federico M. Federici and Sharon O'Brien, 132–151. London: Routledge.

Parra, Albert, Andrew W. Haddad, Mireille Boutin, and Edward J. Delp. 2018. "A Hand-Held Multimedia Translation and Interpretation System with Application to Diet Management." https://arxiv.org/abs/1807.07149.

Pitman, Jeff. 2021. *Google Translate: One Billion Installs, One Billion Stories*. https://blog.google/products/translate/one-billion-installs/

Pituxcoosuvarn, Mondheera, Toru Ishida, Naomi Yamashita, Toshiyuki Takasaki, and Yumiko Mori. 2018. "Machine Translation Usage in a Children's Workshop." *CollabTech 2018: Collaboration Technologies and Social Computing: 10th International Conference, Proceedings*: 59–73.

PRH: Patentti- ja rekisterihallitus [Finnish Patent and Registration Office]. 2018. *Patenttiopas*. Helsinki: Patentti- ja rekisterihallitus (PRH).

Pym, Anthony. 2015. "Translating as Risk Management." *Journal of Pragmatics* 85 (Aug): 67–80.

Pym, Anthony. 2021. "Translation, Risk Management and Cognition." In *The Routledge Handbook of Translation and Cognition*, edited by Fabio Alves and Arnt Lykke Jakobsen, 445–458. Abingdon, Oxon, UK: Taylor and Francis Group.

Pym, Anthony and Kayo Matsushita. 2018. "Risk Mitigation in Translator Decisions." *Across Languages and Cultures* 19 (1): 1–18.

Pym, Anthony, François Grin, Claudio Sfreddo, and Andy L. J. Chan. 2012. *The Status of the Translation Profession in the European Union*: European Commission. https://op.europa.eu/en/publication-detail/-/publication/4e126174-ea20-4338-a349-ea4a73e0d850.

Risku, Hanna. 2014. "Translation Process Research as Interaction Research: From Mental to Socio-Cognitive Processes." *MonTI* (Special Issue 1): 331–353.

Risku, Hanna and Regina Rogl. 2021. "Translation and Situated, Embodied, Distributed, Embedded and Extended Cognition." In *Routledge Handbook of Translation and Cognition*, edited by Fabio Alves and Arnt Lykke Jakobsen, 478–499. Abingdon, Oxon, UK: Routledge.

Risku, Hanna, Regina Rogl, and Jelena Milosevic. 2019. "Introduction: Translation Practice in the Field: Current Research on Socio-Cognitive Processes." In *Translation Practice in the Field: Current Research on Socio-Cognitive Processes*, edited by Hanna Risku, Regina Rogl and Jelena Milosevic, 2–24. Amsterdam: John Benjamins.

Rossetti, Alessandra, Sharon O'Brien, and Patrick Cadwell. 2020. "Comprehension and Trust in Crises: Investigating the Impact of Machine Translation and Post-Editing." *Proceedings of the 22nd Annual Conference of the European Association for Machine Translation*: 9–18.

Rossi, Laura and Dion Wiggins. 2013. "Applicability and Application of Machine Translation Quality Metrics in the Patent Field." *World Patent Information* 35 (2): 115–125.

Roturier, Johann. 2006. "An Investigation into the Impact of Controlled English Rules on the Comprehensibility, Usefulness, and Acceptability of Machine-Translated Technical Documentation for French and German Users." Ph.D. dissertation, Dublin City University.

Saldanha, Gabriela and Sharon O'Brien. 2014. *Research Methodologies in Translation Studies*. Abingdon, Oxon UK: St Jerome Publishing.

Scott, Juliette and John O'Shea. 2021. "How Legal Documents Translated Outside Institutions Affect Lives, Businesses and the Economy." *International Journal for the Semiotics of Law*: 1–43. https://link.springer.com/article/10.1007/s11196-020-09815-5.

Shigenobu, Tomohiro. 2007. "Evaluation and Usability of Back Translation for Intercultural Communication." In *Usability and Internationalization: Global and Local User Interfaces*, edited by Nuray Aykin. Vol. 4560, 259–265. Berlin, Heidelberg: Springer Berlin Heidelberg.

Smith, Ross. 2003. "Overview of PwC/Sytranet On-line MT Facility." *Translating and the Computer 25:* 1–17.

Somers, Harold. 1997. "Machine Translation and Minority Languages." *Translating and the Computer 19*: 1–13.

Somers, Harold and Elizabeth Wild. 2000. "Evaluating Machine Translation: The Cloze Procedure Revisited." *Translating and the Computer 22:* 1–11.

Specia, Lucia and Kashif Shah. 2018. "Machine Translation Quality Estimation: Applications and Future Perspectives." In *Translation Quality Assessment: From Principles to Practice*, edited by Joss Moorkens, Sheila Castilho, Federico Gaspari and Stephen Doherty, 201–235: Springer.

Stewart, Osamuyimen, David Lubensky, Scott Macdonald, and Julie Marcotte. 2010. "Using Machine Translation for the Localization of Electronic Support Content: Evaluating End-User Satisfaction." *AMTA 2010: The Ninth Conference of the Association for Machine Translation in the Americas*.

Suojanen, Tytti, Kaisa Koskinen, and Tiina Tuominen. 2015. *User-Centered Translation*. *Translation Practices Explained,* edited by Sharon O'Brien, Kelly Washbourne. Oxon, UK: Routledge.

Suzuki, Hiroshi and Reiko Hishiyama. 2016. "An Analysis of Expert Knowledge Transmission using Machine Translation Services." *Proceedings of the Seventh Symposium on Information and Communication Technology*: 352–359.

Thicke, Lori. 2011. "Turbocharged MT Testing at Cisco." *MultiLingual*, 21–23.

Thicke, Lori. 2013. "Revolutionizing Customer Support through MT." *MultiLingual*, 49–52.

Tinsley, John. 2017. "Machine Translation and the Challenge of Patents." In *Current Challenges in Patent Information Retrieval*, edited by Mihai Lupu, Katja Mayer, Noriko Kando and Anthony J. Trippe, 409–431. Berlin, Germany: Springer-Verlag GmbH.

Tinsley, John, Alexandru Ceausu, Jian Zhang, Heidi Depraetere, and Joeri Van de Walle. 2012. "IPTranslator: Facilitating Patent Search with Machine Translation." *AMTA-*

2012: Proceedings of the Tenth Biennial Conference of the Association for Machine Translation in the Americas.

Tuominen, Tiina. 2012. "The Art of Accidental Reading and Incidental Listening: An Empirical Study on the Viewing of Subtitled Films." Ph.D. dissertation, University of Tampere.

Turner, Anne M., Yong K. Choi, Kristin Dew, Ming-Tse Tsai, Alyssa L. Bosold, Shuyang Wu, Donahue Smith, and Hendrika Meischke. 2019. "Evaluating the Usefulness of Translation Technologies for Emergency Response Communication: A Scenario-Based Study." *JMIR Public Health and Surveillance* 5 (e11171).

Turovsky, Barak. 2016. *Google Translate Blog: Ten Years of Google Translate*. blog.google/products/translate/ten-years-of-google-translate/

United Nations. 1948. *Universal Declaration of Human Rights*.

USPTO: U.S. Patent and Trademark Office. 2018. *Manual of Patent Examining Procedure* U.S. Patent and Trademark Office.

van der Meer, Jaap. 2006. "The Emergence of FAUT Fully Automated Useful Translation." *Proceedings of the 11th Annual Conference of the European Association for Machine Translation*.

van der Meer, Jaap, Attila Görög, Dace Dzeguze, and David Koot. 2017. *Measuring Translation Quality: From Translation Quality Evaluation to Business Intelligence*. De Rijp, The Netherlands: TAUS BV.

Vasconcellos, Muriel. 1994. "The Current State of MT Usage Or: How do I use Thee? Let Me Count the Ways." *The LISA Forum Newsletter*, 1–8.

Vasiljevs, Andrejs, Rihards Kalnins, Marcis Pinnis, and Raivis Skadins. 2014. "Machine Translation for E-Government – the Baltic Case." *Proceedings of AMTA 2014*: 181–193.

Vieira, Lucas Nunes, Minako O'Hagan, and Carol O'Sullivan. 2020. "Understanding the Societal Impacts of Machine Translation: A Critical Review of the Literature on Medical and Legal use Cases." *Information, Communication & Society*: 1–18.

Villalobos, Oscar, Stephanie Lynch, Conni DeBliek, and Linda Summers. 2017. "Utilization of a Mobile App to Assess Psychiatric Patients with Limited English." *Hispanic Journal of Behavioral Sciences* 39 (3): 369–380.

Voida, Amy, Elizabeth D. Mynatt, Thomas Erickson, and Wendy A. Kellogg. 2004. "Interviewing Over Instant Messaging." *Proceedings of the 17th ACM International Conference on Supporting Group Work*: 1344–1347.

Way, Andy. 2013. "Traditional and Emerging use-Cases for Machine Translation." *Paper presented at Translating and the Computer 35*.

Way, Andy. 2018. "Quality Expectations of Machine Translation." In *Translation Quality Assessment: From Principles to Practice*, edited by Joss Moorkens, Sheila Castilho, Federico Gaspari and Stephen Doherty, 159–178. Cham, Switzerland: Springer.

Wendt, Chris. 2012. "Machine Translation at Microsoft". *Translation and the Computer 34*.

Wendt, Chris and Federico Garcea. 2013. "USE CASE: Customization and Collaboration to Enhance MT for a Knowledge Base Online Portal." *Proceedings of the XIV Machine Translation Summit*.

Williams, Lawrence. 2006. "Web-Based Machine Translation as a Tool for Promoting Electronic Literacy and Language Awareness." *Foreign Language Annals* 39 (4): 565–578.

Xu, Bin, Ge Gao, Susan R. Fussell, and Dan Cosley. 2014. "Improving Machine Translation by Showing Two Outputs." *Chi '14: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*: 3743–3746.

Yamada, Setsuo, Syuuji Kodama, Taeko Matsuoka, Hiroshi Araki, Yoshiaki Murakami, Osamu Takano, and Yoshiyuki Sakamoto. 2005. "A Report on the Machine Translation Market in Japan." *Proceedings of the Tenth Machine Translation Summit*: 55–62.

Yamashita, Naomi and Toru Ishida. 2005. "Analyzing Misconceptions in Multilingual Computer-Mediated Communication." *GROUP '05 Proceedings of the 2005 international ACM SIGGROUP conference on supporting group work*: 352–353.

Yamashita, Naomi and Toru Ishida. 2006a. "Effects of Machine Translation on Collaborative Work." *CSCW '06 Proceedings of the 2006 20th anniversary conference on Computer supported cooperative work*: 515–524.

Yamashita, Naomi and Toru Ishida. 2006b. "Automatic Prediction of Misconceptions in Multilingual Computer-Mediated Communication." *IUI '06: Proceedings of the 11th international conference on Intelligent user interfaces*: 62–69.

Yamashita, Naomi, Rieko Inaba, Hideaki Kuzuoka, and Toru Ishida. 2009. "Difficulties in Establishing Common Ground in Multiparty Groups using Machine Translation." *CHI '09: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*: 679–688.

Yang, Jin and Elke Lange. 1998. "SYSTRAN on AltaVista: A User Study on Real-Time Machine Translation on the Internet." *Proceedings of the 3rd Conference of the Association for Machine Translation in the Americas "Machine Translation and the Information Soup*: 275–285.

Yang, Jin and Elke Lange. 2003. "Going Live on the Internet." In *Computers and Translation: A Translator's Guide*, edited by Harold Somers, 191–210. Amsterdam/Philadelphia: John Benjamins.

Yasouka, Mika and Pernille Bjorn. 2011. "Machine Translation Effects on Communication: What Makes it Difficult to Communicate through Machine Translation?" *Proceedings of the Second International Conference on Culture and Computing*: 110–115.

Yufrizal, Hery. 2001. "Negotiation of Meaning and Language Acquisition by Indonesia EFL Learners." *TEFLIN Journal* XII (1): 60–87.

Zaveckaite, Asta and Aureilija Ulbinaite. 2018. "Assessment Criteria of Project Risk Management in Language Translation Service Companies." *Technological and Economic Development of Economy* 24 (4): 1323–1343.

PUBLICATIONS

Article I — Nurminen, Mary and Niko Papula. 2018. "Gist MT Users: A Snapshot of the Use and Users of One Online MT Tool." *Proceedings of the 21st Annual Conference of the European Association for Machine Translation:* 199–208.

Article II — Nurminen, Mary. 2019. "Decision-making, Risk, and Gist Machine Translation in the Work of Patent Professionals." *Proceedings of the 8th Workshop on Patent and Scientific Literature Translation*: 32–42.

Article III — Nurminen, Mary. 2020. "Raw machine translation use by patent professionals: A case of distributed cognition." *Translation, Cognition and Behavior* 3 (1): 100–121.

Article IV — Nurminen, Mary. 2016. "Machine Translation-Mediated Interviewing as a Method for Gathering Data in Qualitative Research: A Pilot Project." *New Horizons in Translation Research and Education* 4 (16): 66–84. Joensuu, Finland: University of Eastern Finland.

Article V — Nurminen, Mary and Maarit Koponen. 2020. "Machine translation and fair access to information." *Translation Spaces* 9 (1): 150–169.

PUBLICATION I

Gist MT Users: A Snapshot of the Use and Users of One Online MT Tool

Mary Nurminen and Niko Papula

Proceedings of the 21st Annual Conference of the European Association
for Machine Translation: 199–208
http://rua.ua.es/dspace/handle/10045/76049

Gist MT Users: A Snapshot of the Use and Users of One Online MT Tool

Mary Nurminen
University of Tampere
Kalevantie 4
FI-33100 Tampere, Finland
mary.nurminen@uta.fi

Niko Papula
Multilizer
Kimmeltie 3
FI-02110 Espoo, Finland
niko.papula@multilizer.com

Abstract

This study analyzes usage statistics and the results of an end-user survey to compile a snapshot of the current use and users of one online machine translation (MT) tool, Multilizer's PDF Translator[1]. The results reveal that the tool is used predominantly for assimilation purposes and that respondents use MT often. People use the tool to translate texts from different areas of life, including work, study and leisure. Of these, the study area is currently the most prevalent. The results also reveal a tendency for users to machine translate documents that are in languages they have some understanding of, rather than texts they do not understand at all. The findings imply that gist MT is becoming a part of people's everyday lives and that perhaps people use gist MT in a different way than they use publishing-level translations.

1. Introduction

Online machine translation (MT) tools have been in use for almost 25 years and people are finding numerous ways to integrate MT into the processes of their everyday lives. However, although research on professional translators' use of MT has grown rapidly, the literature on all other users of MT remains limited. This paper aims to contribute to that limited body of research with a study on the users of one online MT tool, Multilizer's PDF Translator.

1.1. Purpose of the Study

Our study focuses on users of MT for assimilation, or scenarios in which people use raw, unedited machine translated text for some other purpose than editing it for publication. Because users most often want just a basic understanding of the information (or gist) of the text, we term them *gist MT users.* We also use it because it is shorter than the term *users of MT for assimilation*; however, we use the two terms interchangeably.

The overall purpose of the study is to present a snapshot of the use and users of one online MT tool. Our questions concern who is using MT, where these users are, how they are using it, when they are using it, and in what areas of life they are using it.

We had several motivations in doing the study. First, because online MT is in such wide use today, we can assume that the number of gist users is much larger than the number of professional translator users. Yet the latter group has been studied far more than gist MT users. We believed it was time to put some focus on other user groups and we hoped to contribute to that with this study. Second, our literature review revealed only one gist MT user survey conducted in the past 10 years. We felt it was time to conduct another one. Finally, this analysis will serve as a basis for a second study we are planning, a qualitative study that will probe more deeply into the specific ways people are using gist MT.

1.2. Related Work

The pioneer study of MT users, by Henisz-Dostert in 1979, was also the first study on gist MT users. In the 40 years since it was published, a relatively small number of articles have been written about gist MT users. These studies can be grouped into two categories: experimental studies on *potential* users of gist MT and survey studies on *actual* gist MT users.

In the experimental studies, groups of potential gist MT users were asked to evaluate specific aspects of MT or the use of MT. Fuji et al. (2001) tested user success with machine translat-

[1] pdf.multilizer.com

ed texts, measured through reading comprehension, against users' *impressions* of comprehensibility and awkwardness. Gaspari (2006) had users evaluate their *confidence* in understanding raw MT. Bowker and Ehgoetz (2007), Bowker (2009), Bowker and Buitrago (2015) and Castilhjo and O'Brien (2017) had users evaluate the *acceptability* of raw MT. They often had users compare preference or acceptability of raw MT, post-edited MT, and human translation. Gaspari (2004), Stewart et al. (2010), and Doherty and O'Brien (2012) had users evaluate raw MT against traditional usability criteria. Finally, Doherty and O'Brien (2014) used eye tracking to measure MT output usability.

The studies on actual gist MT users include research on market or usage reports, end-user surveys, or a combination of the two. A small number of these were agnostic to MT systems, focusing on groups who were using any number of the systems available at the time. A larger group of research focuses on users of one specific system. A limitation of this second group of studies is that they describe only a specific type of user and therefore the results cannot be considered representative of all MT users. However, they do contribute information on those users and, seen collectively, help to paint an overall picture of gist MT usage.

The first studies on users of various MT systems were sponsored by the International Association for Machine Translation (IAMT) in 1993 and 1995. These studies used participants they recruited through the manufacturers of MT systems or through the AMTA website. Although they focused mainly on professional translators, who used MT for dissemination, they did include a small amount of data on gist MT users in the form of eight testimonials (Lawson and Vasconcellos, 1993). The Asia-Pacific Association for Machine Translation (AAMT) recruited participants for a series of studies in 2003-2005 through their website, so the user group represented was again not specific to any one tool. These surveys focused much more on gist MT users, indicating that "the main use of machine translation" was assimilation (Yamada et al. 2005, p. 58). The final study that was not dependent on any one MT tool was that carried out by Gaspari in 2007. The survey, conducted at several UK university campuses, used students as informants and covered user demographics, experience with computers and MT, languages translated, use of MT for assimilation vs. dissemination, genres translated, and user evaluations of MT.

The first study that focused on users of a specific system was the study on the users of the Georgetown MT system cited earlier in this article (Henisz-Dostert, 1979). It used a survey, although that survey was administered almost entirely through face-to-face interviews. It provided a rich and multifaceted description of the users, how they used the system, and their experience regarding usefulness, speed, and quality. The study also included a few interesting questions on how users experience cognitive processes, which subsequent surveys have not touched on. These included questions such as "If the style of the MT is awkward, can you correct it mentally?" and "Do you get 'used to' reading MT?" (Henisz-Dostert 1979, p. 193) The only other study we are aware of that address cognitive processes was Doherty and O'Brien's (2014) previously mentioned eye tracking study.

The next study of the users of one system was conducted in Japan by Hoshino (1995), focusing on users of the Korya Eiwa ("It's Nice! English–Japanese") consumer desktop system. The survey was comprehensive, covering user demographics, genres and subject matters translated, users' fluency in English, experience with MT, purpose, motivations, and expectations for MT. Flanagan's (1996) paper described the usage of CompuServe's online MT service as well as users' reactions to it. Another online service, AltaVista Translation with Systran, was the focus of a study by Yang and Lange (2003). The study included both an analysis of usage and feedback data in the form of 5,005 e-mails received in 1998.

A few studies have been conducted on company-internal MT systems and their users. Smith (2003) analyzed PriceWaterhouseCooper's intranet-based MT system and its users. This was perhaps the first study on a system that supports a large number of language pairs, 37 in total. It described how people used the system, their reactions to it, and factors that affected users' satisfaction with the system. Another company-internal study was conducted by Nuutila (2005), who reported on a survey conducted with users of Nokia's Roughlate MT service.

The latest user study we are aware of was a study by Burgett (2015) on the users of Intel's machine-translated support content. This study asked users to perform usability tests while working with Intel's machine translated content.

2. Multilizer's PDF Translator

The tool in our study, PDF Translator, is an online MT tool that translates full documents that are in either PDF or Word format. A user submits a document, then the tool extracts the texts, puts them through machine translation, rebuilds the document with the original pictures in place, and returns it to the user in the requested language. PDF Translator utilizes the MT engines of Microsoft, Google and PROMT to perform the translations. Due to the proprietariness of the engines and the dynamic nature of MT development, we do not have information on the exact type of MT (rule-based, statistical or neural) used for each language pair during the time of the study.

PDF Translator is meant for any type of document that people want to have translated, so it is not trained for specific genres or subject matters. Two versions are available, a desktop and an online version. The desktop version, which was developed first, is downloaded onto the user's computer and used from there. Its user interface is available in 14 languages. Users can translate up to 3 pages at a time for a total of 15 pages for free. PDF Translator offers three levels of paid licenses: Standard, Pro and Business, and after initial purchase of a license, additional pages can be purchased in batches. The desktop version supports 47 source languages and 39 target languages. The newer online version has been in use since 2016 and it is currently available through an English, Spanish or Chinese user interface. Users can translate a small amount of text (one page) free of charge and thereafter they can purchase packages of translation (10, 50, 100, etc. pages). The online version supports translation between 42 languages.

2.1. MT for PDF and DOC Documents

One important aspect of PDF Translator is that it translates entire documents instead of pieces of text typed or copy/pasted into a text field. This holds several implications for our study and the types of users it addresses. First, it excludes incidences when people enter only one or two words, essentially using MT as a bilingual dictionary. Previous studies have found this to constitute a large portion of MT use. For example, Yang and Lange reported that "more than 50% of translations are of one- or two-word phrases" (Yang and Lange, 2003, p. 199) and Gaspari was led to devote a whole section of his PhD to "(Mis-)Using Free Web-based MT Services as Online Dictionaries" (Gaspari, 2007, p. 108). Another implication of translating whole documents is that the materials people submit for translation tend to be well-formed and written, published documents instead of more informal texts such as chat messages or personal correspondence. This can influence the areas of life where people use MT – for work and study or in their free time. A final implication is that, due to the very nature of PDF as a publication instead of an editing format, users are far more likely to be gist MT users than to be people who want to edit the material for publication. All of these factors contribute to profiling a specific type of user and need to be kept in mind when reading this study.

3. Materials and Methods

Our goal was to capture a snapshot of the use and users of PDF Translator in a short, specific point of time. We chose a four-month period, November 1, 2017 through February 28, 2018, and collected two types of data from the period for analysis. We collected log files from both the desktop and the online systems, and we conducted an online end-user survey with users of the desktop system.

Our first batch of data consisted of the logs from the desktop and online versions of PDF Translator. We used the logs to examine the times that submissions for translation were made, the places they were made from, and the source and target languages involved.

The end-user survey was short, consisting of eight questions in three categories:

Category	Questions
Basic demographics	1. What is your gender? 2. What is your age? 3. What language are you most proficient in? 4. What is the highest degree or level of school you have completed?
Frequency of use of MT tools	5. How often do you use tools that automatically translate texts, similarly to PDF Translator or Google Translate?
Questions on the specific document submitted for translation	6. Why did you want to translate the document? 7. Did you need the document for work, study, or leisure purposes? 8. How well did you understand the language of the original written document?

Table 1. Survey questions.

The reason for the brevity of the survey was that, in keeping with the idea of a snapshot, our focus was on quantity more than quality. The survey needed to be short enough so that a large number of people would be willing to answer it.

Besides keeping the survey short, we used other strategies to encourage users to respond. We offered all respondents the chance to participate in a drawing for five small prizes: 100 pages of free translation through PDF Translator. We also named it *3-minute Survey for Users of PDF Translator* under the assumption that precise information on how long it would take to answer the survey would encourage people decide to devote time to it. The average response time was, in fact, three minutes.

Due to limited resourcing, we had to make decisions on what languages to offer the survey in. We decided to offer the survey to users of the most popular 6 of the 14 languages the desktop version of PDF Translator is available in: English, Spanish, Portuguese, French, Russian and Indonesian.

An invitation to answer the survey was offered to users after they had submitted a document into PDF Translator and received the translation back. It was offered to everyone who submitted a document during that period, meaning that both heavy users of the tool and first-timers could answer.

4. Discussion

Besides the log files, our data included 1,579 responses to the three-minute survey. The response distribution by language survey is displayed in the following table.

Language survey	Number of responses
Spanish	652
Portuguese	283
French	211
Russian	188
English	147
Indonesian	98
Total	**1579**

Table 2. Survey response distribution.

PDF Translator has a large customer base in Spanish-speaking countries and this is reflected in the high number of responses to the Spanish survey. The placement of the other language surveys correlate roughly with our statistics on the countries and target languages with the most traffic during the study period. While compiling responses, we noticed that a large number of responses to the English survey (49 responses, comprising 25% of all responses), were from people who marked Indonesian as their most proficient language. We did not observe a similar phenomenon in any other language survey. We decided to move these 49 responses from the English survey to the Indonesian one. The previous table reflects the numbers *after* that change.

4.1. Locations and Languages

PDF Translator is used widely across the world. Our logs indicated that requests for translation during the study period came from 181 countries and territories. The tool's large customer base in Spanish-speaking countries is reflected in the list of the countries with the most traffic, with 10 of the top 20 spots being occupied by those countries. Other countries in the top 20 include Brazil, Indonesia, Poland, Germany, Italy, Russia, Turkey, France, Ukraine, and Portugal.

English was the most popular source language, with 85% of all documents translated during the study period being originally in English. The next languages on the list of source languages included German, Spanish, French, Portuguese, Italian, Russian, Polish, Dutch and Indonesian. Spanish led the list of the most popular target languages, followed by Portuguese, English, French, Russian, Indonesian, German, Polish, Italian and Turkish.

The top language pair of English–Spanish comprised 47% of all requests. This was expected, considering PDF Translator's customer base. Also, this language pair has appeared at the top of lists in survey and market studies for a long time, including those by Yang and Lange (2003), Smith (2003), Gaspari and Hutchins (2007) and Turovsky (2016).

Indonesian's position near the top of the language lists was interesting. The past ten years have seen a major expansion in the language palette of online MT tools (e.g. Turovsky, 2016). It appears that this expansion is beginning to produce results and new language pairs are emerging at the top of the lists of the most-translated languages. For example, Google's recent reports on the most-translated languages include the ones that have appeared at the top of these lists for years—Spanish, Russian and Portuguese—but also relative newcomers to online MT, such as Arabic and Indonesian (Turovsky, 2016). Indonesian proved to be an interesting and different market in other areas of our study as well.

4.2. Survey Participant Demographics

The overall gender demographic of survey participants showed males comprising 68% of responses, females 32%, and the group of *other*, 3%. Small differences surfaced when comparing the results of different language areas. In the Portuguese, Spanish and English surveys, males made up 61–68% of responses while in the French and Russian surveys, 82–83% of respondents were male. Indonesia was the only country in which female respondents outnumbered male (54% and 46%, respectively). The high proportion of men in most of the language surveys seems to be typical in studies of technological systems.

The age distribution shown in survey answers was also typical of that shown in technology studies, with the 19–29 age group providing the largest number of responses, 46% altogether. Similarly to the results of the gender demographic, the age demographic also contained differences in the results from different language surveys, as is shown in the following figure.

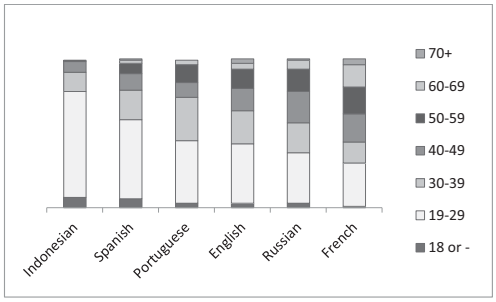


Figure 1. Age distribution of respondents in different language surveys.

Indonesian again displayed a different profile from the other surveys. In that survey, the 19-29 age group made up 71% of the total, 18 percentage points higher than the next (Spanish) survey. The French and Russian surveys were again at the opposite end of the scale, with a much more even distribution of ages. Another interesting point was that the French-speaking older respondents seem to be the most active. Whereas in most of the language surveys, the two highest age groups comprised 3–7% of respondents, in the French survey this group comprised 19% of all respondents. Although the total overall number of answers in the highest age groups, 60–69 and 70 or older, was small (68 and 19 responses respectively), it was good to note that people in these age groups are also using MT actively.

The following figure shows how much of each respondent age group was comprised of female, male and other genders.

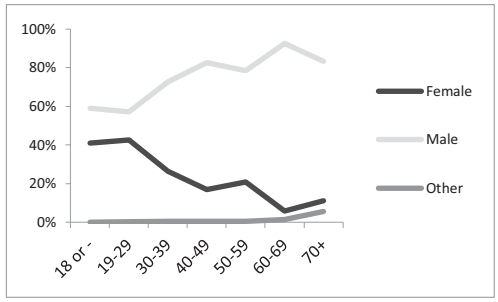


Figure 2. Percentage distribution of survey respondents by gender and age group.

The chart shows that in the younger age groups, a smaller gap exists between the male and female composition of the respondents. This gap grows and peaks in the 60–69 age group before becoming smaller again in the 70 or older group. A somewhat even number of people identify as some other gender throughout all age groups, although the relatively small overall number of respondents in the 70 or older group resulted in the *other* group comprising a higher percentage of the whole.

The highest degree or level of school reported by respondents is shown in the following table.

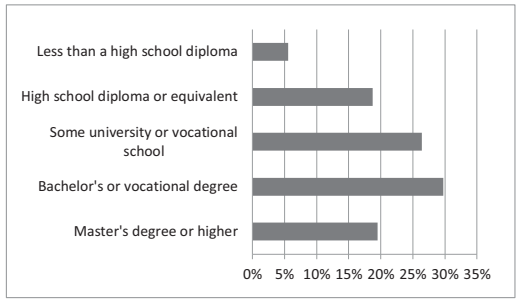


Figure 3. Highest level of education of respondents.

Respondents appear to be fairly highly educated, with the largest group being comprised of people who already have a vocational or bachelor's degree. In comparing the different language surveys, the French and Russian surveys once again stood out in that they had high percentages of respondents who held a master's degree or higher. In fact, the educational level with the most responses in both surveys was a master's degree.

4.3. Frequency of MT Use

As has been noted by Gaspari (2007) and others, a self-administered survey such as this one can result in responses being given by people who are relatively more active in the technology area than the general user population. This factor needs to be considered when examining the responses to our survey question on how often respondents use MT, which are displayed in the following chart.

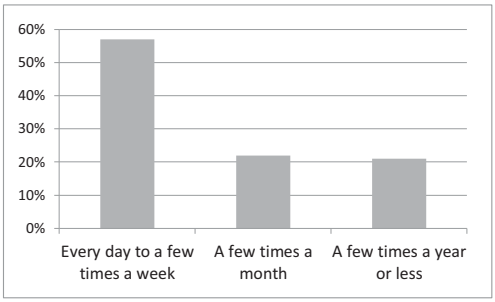


Figure 4. How often respondents report using machine translation.

These results indicate that a majority of the overall respondents of this survey tend to use MT on a very regular basis. In comparing to previous studies that have asked this question, Yamada et al. (2005) reported that only 13–18% of users used MT as frequently in 2003–2005. However, Nuutila's (2005) study showed that 63% of Nokia's in-house Roughlate system users reported using the system several times a day or at least every week.

The next chart shows a breakdown of reported frequency of use by age group.

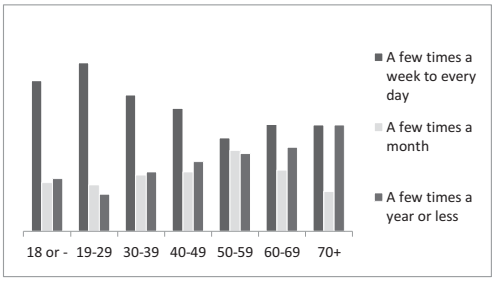


Figure 5. Frequency of MT use by different age groups.

As is shown here, the younger age groups, 18 or under and 19–29, showed a stronger tendency to use MT very frequently than respondents in older age groups. In fact, the level of very frequent use for the 19–29 group was remarkably high, 67%.

4.4. Purpose: Assimilation, Dissemination, or Something Else

To explore users' purposes for using MT, the area of life they were using MT in, and their proficiency in the languages involved, the survey included three questions that asked specifically about the document the respondent had submitted for translation right before being invited to take the survey. The first of these questions concerned whether users were using the submitted document for assimilation, dissemination, or some other purpose. Although we could assume that people translating whole documents (many of them PDFs) are mainly using MT for assimilation purposes, we wanted to verify this. We started with the questions and answer choices used by Gaspari in his survey of students (Gaspari, 2007, p. 102–103) and edited them a bit. The following table shows the overall responses.

Why did you translate the document?	**% of responses**
I wanted to understand it myself. (assimilation)	58%
I wanted to verify that I understood it myself. (assimilation)	18%
I wanted to translate it into my own language so that someone else can understand it. (assimilation for other person)	14%
I wanted to translate it from my language into another language so that someone else can understand it. (dissemination)	6%
Some other reason (please specify).	4%

Table 3. Purpose of translating the document submitted for translation.

Combining the first and second answers gives an overall view to assimilation and shows that a majority of respondents, 76%, are indeed using the machine-translated documents for their own assimilation. However, the second answer taken alone is also interesting in that it shows that people are using MT for understanding documents, but also for verifying their understanding. Another interesting point arises when comparing the responses of different language surveys. In Indonesia, 25% of respondents reported that they translated the document into their own language so that someone else could understand it. In other language surveys, the rate was only 10–16%. Combining this with the relatively young de-

mographics of that market, could this reflect an effort by younger people to help their technologically more reticent elders?

4.5. Area of Life Where MT was Used

The second of the questions we asked about the document the respondent had translated regarded the area of life that the document concerned: work, study, or leisure. We allowed respondents to select more than one choice in case the document was used in various areas. However, only 11% chose more than one area. The following figure displays the overall compiled results of responses to the question.

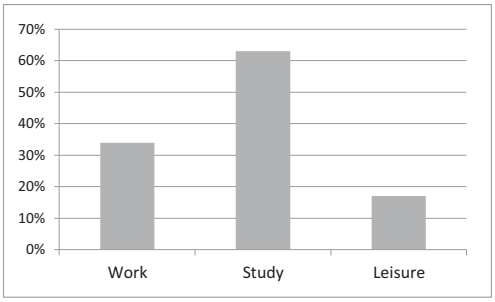


Figure 6. Percentage of respondents who listed work, study, and/or leisure as the purpose of the document they translated.

Overall, 63% of the respondents reported that at least one of the areas of life in which they needed the translated document was study. This would indicate that, at least for the type of user who is translating whole documents (and willing to answer surveys), MT is being used widely for learning purposes.

This figure shows the responses by age group.

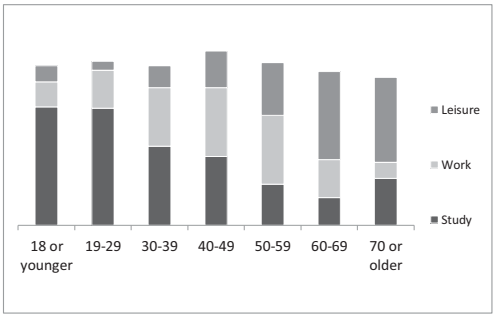


Figure 7. Reported area of life where machine translated document was used, by age group.

This distribution seems logical and perhaps expected, with users in the younger age groups showing a relatively strong emphasis on study. It is interesting that the study category increased again in the 70 or older age group, though it should be kept in mind that the number of responses in that group was small (19), and that respondents who are active users of MT, and are willing to answer surveys, might well also have a keen interest in self-study.

Two factors seem to have contributed to making *study* the top area reported. First, a relatively high number of responses to the survey came from the 19–29 age group. Second, responses from the Spanish and Portuguese surveys were also relatively high, and as can be seen in the following table, both of those languages showed very high scores for *study*.

Survey	Work	Study	Leisure
Indonesian	19%	88%	4%
Portuguese	30%	73%	15%
Spanish	31%	75%	9%
English	46%	49%	19%
French	43%	34%	39%
Russian	44%	36%	31%

Table 4. Percentage of respondents who listed work, study, and/or leisure as the purpose of the document they translated in different surveys.

In this table, the English, French, and Russian answers reflect more of an emphasis on work. In fact, in the French and Russian results, work surpasses study as the area of life the translated document concerned. As discussed earlier in this article, the demographics of the French and Russian respondents were somewhat different than those of the other language surveys. These differences seem to indicate that the way MT is used can be different in different groups or geographical areas.

In addition to analyzing the responses to our survey, we also used the log files to analyze the day of the week and time of day when people requested translations. We converted all log time stamps to local times. The results of that analysis are presented in the following figure, which shows usage levels for the seven days of the week and hour-by-hour. Each of the seven lines in the graph represents one day of the week. Black lines were used for Monday–Thursday and gray for Friday–Sunday.

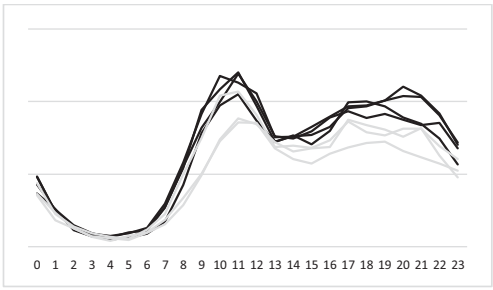


Figure 8. Usage by day of the week and time of day. The black lines are Monday–Thursday and gray lines Friday–Sunday.

Although all lines demonstrate activity during the evening hours, a clearly higher activity level emerges on Monday–Thursday than on Friday–Sunday. This analysis seemed to support the result that study and work are areas of life where users of the tool request translations, more than leisure.

It should be noted that these results reflect the situation for one tool at a specific point in time. As the technology and users mature, the overall emphasis could shift from study to other areas of life. Another point of consideration is that our results do not provide details on the level of education users are at when they use MT for study. It could be anything from grade school through Ph.D. research. The results also do not tell us exactly how users are using the machine-translated information: to help them in language production, for self-study, or to read scientific articles in a language they do not know. These questions should be addressed in future studies.

4.6. Understanding of Source Language

The third question in the survey related to the document that each respondent had submitted for translation was the following: *How well do you understand the language of the original written document (before it was translated)?* The possible answers were *Very well, Well, A little* and *Not at all.*

In the overall results, 51% of people reported that they understood *a little* of the source text and 33% said they understood the source text *well* or *very well*. By contrast, only 17% labeled their understanding as *not at all*. A few differences emerged when comparing the results of different language surveys. The Portuguese and English surveys had the highest percentage of people answering that their understanding of the source language was *not at all* (23% in English, 36% in Portuguese). In all other languages, 15% or fewer reported having no understanding.

As participants reported using PDF Translator for a variety of purposes, including dissemination, we conducted a separate analysis of people who specifically used it for assimilation, or gist users. For that analysis, we used only the answers of respondents who said their reason for translating the text was either that they wanted to understand it themselves or that they wanted to verify that they understood it themselves. As is shown in the following chart, a large majority of this specific group displayed at least a basic understanding of the source texts they translated. This result was similar to the overall results.

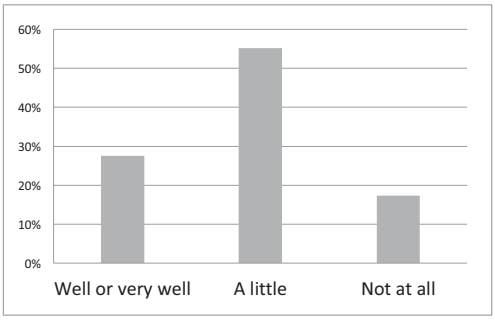


Figure 9. Reported understanding of the source text of the document submitted for translation by gist users.

The responses showed that in general, users of this tool often seem to translate texts that are in languages in which they already have some proficiency. Some previous survey studies have asked about users' competence in the source language, including Henisz-Dostert (1979), Hoshino (1995) and Yamada et al. (2005). A few other studies have uncovered indications of a link between knowledge of the source language and use of MT (Nurminen, 2016; Ogura et al., 2004).

Of course, people who are translating documents they already have some understanding of might also be simply testing PDF Translator or MT. Although we offered such people the choice to answer that their reason for translating the document was *some other reason*, some people may have instead indicated that their purpose was their own understanding and are therefore included in the assimilation group. In spite of this, there did appear to be a tendency to translate documents that respondents already had some understanding of, and this tendency has some interesting implications. First, this might be one

reason why, even with the onslaught of new language pairs available in online MT tools, the same European-based languages still tend to dominate the lists of the most translated languages. Because they are being taught widely in schools, these are languages in which people may have low-to-medium (although existing) competence.

Second, this could reflect a tendency to use MT with caution. Users want to be able to compare the machine translated text to the original so that they can evaluate the general level of MT output. This tendency might decrease in the future, as MT improves and users' trust in its quality increases.

Third, this raises a question that was asked in Henisz-Dostert's survey (1979): how do people find the texts they have machine translated? Do they need to have a basic understanding of the text (or even the title) to be able to make the decision to machine translate it? This would restrict the texts and the languages involved in gist MT use.

Finally, the phenomenon raises a question about how people use MT. Is MT in these cases being used as some type of language tool, which users can combine with other resources, such as their limited competence in the source language or their familiarity with the topic of the text, to gain understanding of a text in another language? If so, does this mean that the way people use gist MT (in raw or possibly also lightly post-edited form) is inherently different than the way they use publishing-level translations? Perhaps we need to begin seeing gist MT as a different translatorial activity than human translation, and to stop comparing them to each other.

5. Conclusions

This study provided a snapshot of the use and users of a specific type of gist MT tool. It presented a picture of who is using PDF Translator, where these users are, how they are using it, when they are using it, and in what areas of life they are using it.

The study confirmed some findings of previous studies. English continues to be the most-translated language and English-Spanish the most commonly translated language pair. However, it also showed that new languages such as Indonesian are beginning to appear at the top of lists of languages involved in MT. The demographics of the survey respondents indicate that, even though overall statistics reflect a bias toward young and male users, which is commonly found in technology studies, differences do emerge in the demographics of different language areas.

A few new tendencies that deserve further study surfaced also. First, gist MT users who translate whole documents seem to use MT often, multiple times a week. Second, the importance of MT in the area of study, at least for the current users of PDF Translator, was a noteworthy result. Finally, users' tendency to machine translate texts in a language that they have some level of proficiency in was a new finding.

Our study shares a limitation with a number of similar surveys in that it studied the users of only one tool and therefore cannot be considered representative of any larger or more general population of users. A second limitation was the use of a self-administered survey, which can lead to a disproportionately enthusiastic picture of MT users. A more random sampling of respondents could produce different results.

The study nevertheless contributes to the small body of literature on gist MT users. The main contribution is that that users' competence in the source language seems to play some role in their use of MT. Users' reports on having some level of proficiency in the source language of the document they translated, plus the tendency some users have to use MT not only for assimilation but also for verifying their understanding of documents, lead to questions of exactly how people are using gist MT. Is it comparable to their use of human translation, or do they use MT in very different ways?

Further studies on how people are using MT in their studies would be called for. We would also like to see new studies that focus on general populations of gist MT users, instead of the users of one tool. However, the most urgent need we envision right now is for deep, qualitative data on exactly how people use gist MT. After the first study in 1979, very little insight has been gained as to how people have integrated MT into their daily lives, what types of processes they use, and the cognitive processes they rely on to extract meaning from imperfect language. As the quality of MT improves and more uses are found for MT in its raw form, the already-pressing importance of this type of data will increase.

References

Bowker, Lynne, and Jairo Buitrago Ciro. 2015. Investigating the usefulness of machine translation for newcomers at the public library. *Translation and Interpreting Studies* 10 (2): 165-186.

Bowker, Lynne, and Melissa Ehgoetz. 2007. Exploring user acceptance of machine translation output: A recipient evaluation. In *Across boundaries: International perspectives on translation studies.* Eds. Dorothy Kenny & Kyongjoo Ryou, 209-224. Newcastle, UK: Cambridge Scholars Publishing.

Bowker, Lynne. 2009. Can machine translation meet the needs of official language minority communities in Canada? A recipient evaluation. *Linguistica Antverpiensia* 8: 123-55.

Castilho, Sheila, and Sharon O'Brien. 2017. Acceptability of machine-translated content: A multilanguage evaluation by translators and end-users. *Linguistica Antverpiensia* 16: 120-136.

Burgett, Will. 2015. Unmoderated remote usability testing of machine translation content. *TAUS Review of Language Business and Technology* IV.

Doherty, Stephen, and Sharon O'Brien. 2014. Assessing the usability of raw machine translated output: A user-centered study using eye tracking. *Intl. Journal of Human-Computer Interaction* 30: 40-51.

—— 2012. A user-based usability assessment of raw machine translated technical instructions. 10th Conference of the Association for Machine Translation in the Americas, San Diego, California, USA.

Fuji, M., N. Hatanaka, E. Ito, S. Kamei, H. Kumai, T. Sukehiro, T. Yoshimi, and H. Isahara. 2001. Evaluation method for determining groups of users who find MT “useful”. MT Summit VIII, 2001, Santiago de Compostela, Spain.

Flanagan, Mary. 1996. Two years online: Experiences, challenges and trends. Proceedings of the 2nd Conference of the Association for Machine Translation in the Americas "Expanding MT Horizons", Montreal, Canada.

Gaspari, Federico. 2007. The role of Online MT in Webpage Translation. Ph.D. thesis, University of Manchester.

——2006. The added value of free online MT services: Confidence boosters for linguistically-challenged internet users, a case study for the language pair Italian-English. 7th Conference of the Association for Machine Translation of the Americas, Cambridge, Massachusetts, USA.

—— 2004. Online MT services and real users' needs: An empirical usability evaluation. 6th Conference of the Association for Machine Translation in the Americas, AMTA 2014, Washington, DC, USA

Henisz-Dostert, Bozena. 1979. Users' evaluation of machine translation. *Machine Translation,* ed. Werner Winter, 149-244. The Hague: Mouton Publishers.

Hoshino, Sadao. 1995. Survey report: Korya eiwa: 100-dollar commodity has expanded the market. *Newsletter of the International Association for Machine Translation* (12).

Lawson, Veronica, and Muriel Vasconcellos. 1993. Forty ways to skin a cat: Users report on machine translation. Machine Translation Today: Translating and the Computer 15, London, UK.

Nuutila, Pertti. 2005. Rough Machine Translation in the Communication Process. Licentiate thesis, University of Tampere.

Nurminen, Mary. 2016. Machine translation-mediated interviewing in qualitative research: A pilot project. *New Horizons in Translation Research and Education* 4.

Ogura, Kentaro, Yoshihiko Hayashi, Saeko Nomura, and Toru Ishida. 2004. User adaptation in MT-mediated communication. The 1st International Joint Conference on Natural Language Processing (IJCNLP-04), Hainan Island, China.

Smith, Ross. 2003. Overview of PwC/Sytranet on-line MT facility. The 25th International Conference on Translating and the Computer, London.

Stewart, Osamuyimen, David Lubensky, Scott Macdonald, and Julie Marcotte. 2010. Using machine translation for the localization of electronic support content: Evaluating end-user satisfaction. Denver, CO, USA.

Turovsky, Barak. April 28, 2016. *Ten years of Google Translate* [Blog post]. Retrieved from https://blog.google/products/translate/ten-years-of-google-translate/.

Yamada, Setsuo, Syuuji Kodama, Taeko Matsuoka, Hiroshi Araki, Yoshiaki Murakami, Osamu Takano, and Yoshiyuki Sakamoto. 2005. A report on the machine translation market in Japan. The 10th Machine Translation Summit, Phuket, Thailand.

Yang, Jin, and Elke Lange. 2003. Going live on the internet. In *Computers and translation: A translator's guide,* ed. Harold Somers, 191-210. Amsterdam/Philadelphia: John Benjamins.

PUBLICATION II

Decision-making, Risk, and Gist Machine Translation in the Work of Patent Professionals

Mary Nurminen

Proceedings of the 8th Workshop on Patent and Scientific Literature Translation: 32–42
https://trepo.tuni.fi/handle/10024/122130

Decision-making, Risk, and Gist Machine Translation in the Work of Patent Professionals

Mary Nurminen
Tampere University
Kalevantie 4
FI-33100 Tampere, Finland
mary.nurminen@tuni.fi

Abstract

This is the first study on how patent professionals use gist machine translation (MT) in their work. Inductive, qualitative research methods were adopted to explore the role of gist MT specifically in decision-making. Results show that certain decisions by patent professionals rely on gist MT, that the decision to involve human translation is often based on a risk assessment, and that certain factors in the patent environment give affordances for the use of gist MT. The study contributes to the body of knowledge on patent MT users and on gist MT users in general.

1 Introduction

Machine translation (MT) for patents has been developed for a few decades and a broad body of research is devoted to the technologies and techniques for producing patent MT. The professionals who work with patents – patent attorneys, counsels, examiners, etc. – use this MT in its raw, unedited form to obtain a basic understanding, or gist, of patent documents that they need but that are in languages they do not understand. Although their use of this raw MT (termed *gist MT* in this article) has been widespread for approximately a decade, very little research has been conducted on these MT users. In fact, while the number of studies on one group of professionals who use MT in their work, translators, has increased in recent years, research on other professional groups who use the technology remains scarce.

The main objective of this article is to provide the first study focused specifically on the users of patent MT. The article presents the results of a qualitative, exploratory study based on interviews with a small group of patent professionals who use MT in their daily work. Three themes were investigated for the article: the types of decisions patent professionals make based on machine-translated information, the risk assessment they use when deciding between relying on gist MT or opting for human translation, and finally, the environmental factors that appear to give affordances for the use of gist MT in this context.

Two important aspects of patent MT are not in the scope of this study. First, the article does not focus on the issue of quality of MT output, as that has already been studied in numerous other articles. Instead, I wanted to concentrate on exploring other factors that influence gist MT use. Second, another key application of patent MT is its use by professional post-editors to enhance their translation process. These users are not included in the scope of this study.

The article will help to inform research and solution development in the patent MT field. It will also contribute to studies of different professions' use of gist MT and to a general understanding of gist MT users. Better knowledge of how MT is used in different contexts and what contributes to successful use will help us to define what makes a potential use case good, or conversely poor, for gist MT use. In addition, research on experienced users of this form of artificial intelligence can give us insights into the needs of users of other AI technologies.

The structure of the article is as follows: the next section contains a review of related work. This is followed by a description of the data and methods used in this study. Section 4 discusses the types of decisions that patent professionals report making based on gist MT. Section 5 describes the risk assessments that informants appeared to undertake when deciding on ordering human translation. Section 6 focuses on the factors in the work environment that appear to support the use of gist MT. Final conclusions and suggestions for further studies are presented in Section 7.

2 Related work

To the best of my knowledge, thus far no studies have been conducted on how patent professionals use gist MT in their everyday work. A few experimental studies have been done. Larroyed (2018) and Tinsley et al. (2012) describe evaluation experiments in which one evaluation is performed by real patent professionals. A number of studies describe technical solutions for patent MT, and some of those include discussions of some aspects of MT in patent professionals' work, for example, Tinsley (2017), Rossi and Wiggins (2013), and List (2012). In addition to these, a few studies that focus on patent searchers also allude briefly to MT in patent search, including Joho et al. (2010) and McDonald-Maier (2009).

To date there is only a small body of research on professional areas where gist MT is used. Professional translation has been studied to some extent, though in that industry MT is predominantly used for dissemination and not for gisting. Industries with reported use of gist MT include customer support, academia, medicine and the legal field. Customer support groups began to offer multilingual access to knowledge base articles through gist MT in the early 2000s. However, although several articles describe these solutions (e.g. Stewart et al., 2010; Dillinger and Gerber, 2009), very little user experience research has been undertaken, as stated in one of the few studies on actual users (Burgett, 2015: 30). A growing body of research focuses on the use of MT in academia. Much of this focuses on the effects of MT on education and students, but some of the studies also cover educators' viewpoints, such as Bowker and Eghoetz's (2007) study on the acceptability of MT in a university setting and Bowker and Buitrago's (2019) book on using MT in research. Health care is another field where gist MT is beginning to be researched. Liu and Watts (2019) give a good overview of current studies on mobile MT use in health care. Most recently, John Tinsley describes the emergence of new use cases for gist MT in two different industries: legal and life sciences (Beninatto and Stevens, 2019). Both cases are similar to the patent case in that MT is mainly used to sift through large numbers of documents to categorize and then locate the ones that need further scrutiny and possibly human translation.

Work in the area of risk and translation has examined risk assessment and management strategies either as part of the individual translator's work (Pym, 2015; Pym and Matsushita, 2018), or from the perspective of the translation process and service provider (Canfora and Ottmann, 2018). Canfora and Ottmann (2016) present a model for risk management for internal translation processes, including a risk matrix combining the probability of risk and the potential consequences. A recent paper by Nitzke, Hansen-Schirra and Canfora (2019) introduces a model for assessing the risk associated with using post-edited or gist MT. Nonetheless, the focus of that study is primarily the post-editing context, while scenarios involving unedited MT remain mostly unexplored.

3 Methods

The main data for this study was gathered through interviews with nine patent professionals working in Scandinavia. The term *patent professional* in this study refers to professionals in the intellectual property rights (IPR) field who use their expertise in patents to assist and guide others (internal or external groups) in their IPR processes. These professionals hold a variety of titles, such as Patent Counsel, Patent Attorney, and Patent Examiner. The informants for the study are presented in Table 1.

Type of informant	**N**
Patent professionals working in companies that are active in filing and prosecuting patent applications	5
Patent professionals working in an IPR service provider	2
Patent professionals working in a government patent office	2
Total	**9**

Table 1. Informants interviewed for the study

I included informants from the key areas where patent professionals work: private companies, IPR service providers, and governmental organizations. The largest group consisted of professionals working in companies that file patent applications. This is somewhat reflective of

the 2010 survey by Joho et al., in which 88% of respondents reported working predominately with internal clients (Joho et al., 2010: 16), which indicates that they worked in patent-filing and prosecuting companies. In addition to the interviews, I gathered background information through talking to people involved in creating and maintaining patent MT solutions.

The average age of the informants was 47 and the average length of experience in the IP field was considerable, 17 years. The group was highly educated; all had at least a master's level education and four of the nine held a PhD degree. This is similar to the educational levels reported in Joho et al. (2010).

The interviews were all semi-structured discussions that occurred either at informants' workplaces or through Skype audio calls. They were conducted in the time frame of April 2018 to February 2019. The two themes of context and transparency were explored in the interviews. I used a variety of sources in the development of the questions. ISO standard 9241-210:2010 (ISO, 2010) defines *context of use* through the broad categories of users, goals and tasks of users, and environment, and this was a good starting point. I relied on descriptions of patent processes in official documents (PRH, 2018; EPO, 2018) and other sources (Alberts et al., 2017; Oesch et al., 2014; Joho et al., 2010) to identify the touchpoints users might have with MT and to develop questions around those touchpoints. The questions also developed somewhat over the course of the data-gathering phase.

Most of the interviews were recorded, transcribed with the aid of automatic transcription tools, and then post-edited. One interview was not recorded due to technical difficulties, so the data from that interview consisted of my notes taken during the interview. A total of 12 hours of interviewing was conducted, and 229 pages of transcription and note data compiled for analysis.

The data was analyzed by closely following the thematic analysis method outlined by Braun and Clarke (2013, 2006, n.d.) with additional guidance from Merriam and Tisdell (2016). The data was approached from a semantic perspective, wherein "coding and theme development reflect the explicit content of the data" (Braun and Clarke, n.d.) rather than searching for underlying meanings in the data. One reason for this was that the topic of the use of technology at work was fairly straightforward. Also, the focus was on the context, as described by informants, instead of each informant's personal experiences.

At a point later in the analysis process, a summary of findings was compiled and a member check performed by three of the informants. They were asked to compare the results against their own experiences and to comment on any incongruences they may detect. These comments were then reviewed and incorporated into the analysis.

A qualitative method was chosen for this study for specific reasons. First, it was necessary because this is the first study on how this group uses gist MT, and research at such an early stage often requires inductive, exploratory methods. When designing the study, there simply was not enough knowledge on these users to allow for the crafting of a quantitative study such as a survey. A second reason was that the small body of research on gist MT users in general tends to rely on surveys and laboratory experiments. I believed there was a need for in-depth studies that would give us a more nuanced view of the use of gist MT. I selected interviewing for data-gathering because it proved difficult to persuade exceedingly busy patent professionals to participate in a study using more time-consuming qualitative methods such as diaries. The interviews required a commitment of only 1.5 hours, which seemed to be more tenable.

4 Decisions that rely on gist MT

Rossi and Wiggins (2013: 116) argue that "In the patent field, MT is used as a support tool for performing novelty, validity, infringement or state-of-the-art searches, and to provide a first understanding of the content of retrieved publications." However, is gaining a "first understanding" really the only way patent professionals use MT, or do they actually make decisions based on gist MT? For example, Henisz-Dostert's study of scientists' use of MT to understand Russian scientific articles reported that, contrary to predictions by early scholars that MT would be used only for scanning, scientists used MT "more as a tool of information than as a tool for the selection of information." (Henisz-Dostert, 1979: 180). One goal of this study was to explore the ways in which patent professionals use gist MT and the decisions they make with its help.

4.1 Relevance

One of the primary uses for patent MT, as defined by Tinsley (2017: 411) is "[to] provide an on-demand 'gist' translation of foreign patents for information purposes to determine relevance." The primacy of using gist MT for this purpose was

also found in this study. Informants described how they used gist MT when searching for 'prior art' (patent documents that show that an invention is not new and therefore present an obstacle to patenting it). For each patent document (either a patent or a patent application) found in their search, they need to decide whether it is relevant to the IPR process they are working on or not. Informants discussed four main IPR processes in which they use machine-translated information to determine relevance: (1) the patenting process (Does this invention show enough novelty that it could be patented?), (2) freedom to operate (Can we launch our products in this market or are there patents that we would be infringing on if we launched?), (3) monitoring (Is this patent application sufficiently relevant to our work that I should monitor its progress?), and (4) infringement (Is this patent infringing on one of our patents or are we in danger of infringing on someone else's patent?).

The results of this study reveal that the decision on relevance is very often made without the help of human translation. Therefore, the first decision made is not whether or not a patent document should be sent for human translation, but whether or not it appears to be relevant, and that is determined largely on the basis of gist MT:

> *I would say it's* [the use of MT] *successful in 90 percent of the time because the conclusion is, this is not relevant…So rejecting things from further analysis I think is done 9 out of 10 reviews of the machine translated documents.* (PP4) [1]

It is important to note that the decision on relevance is not as minor a decision as it may seem. The consequences of mistakenly discarding a patent document that seems irrelevant can be considerable, as was reported by informants:

> *…for example I can decide about a patent that it is not in any way relevant for us, which is a pretty strong decision, because then we shut it out completely, the whole followup of the patent, and we just think that that won't be harmful, but then it could be that if there's a mistake in the translation then it turns out that it really* ***is*** *harmful. But those are the kinds of decisions I make.* (PP1) [2]

> *At work we talk about how most mistakes take place because someone* ***overlooks*** *a relevant patent…when a mistake happens, it is most likely to be caused by that. But mistakes can come from other reasons than the machine translation. There are just so many patents to go through. But putting it into the 'not interesting' pile is a risk.* (PP3)

4.2 Monitoring

A second type of decision that is very often made based solely on gist MT is the decision to tag a patent application for monitoring. If an application is deemed relevant, patent professionals may decide to follow its progress throughout its prosecution. Besides being used to determine enough relevance for monitoring, Gist MT is also used to understand communications on the application's prosecution or to review changes in the application. Tagging an application for monitoring also often means that the decision on human translation is postponed, because the application will most likely change before it is granted:

> *…if it's about pending patents then the claim scope is changing all the time, so therefore even if you would translate it and get it kind of right in the beginning, it's something different when it's granted…so therefore there's no point maybe to get it human translated at the early stage* (PP2)

4.3 Patenting and opposition

A third area in which informants reported using gist MT in decision-making was during the patenting process. Within the European context, the role of MT in the examination process is explained in official guidelines:

> *In order to overcome the language barrier constituted by a document in an unfamiliar non-official language, it might be appropriate for the examiner to rely on a machine translation of said document…A translation has to serve the purpose of rendering the meaning of the text in a familiar language…Therefore mere grammatical or syntactical errors which have no impact on the possibility of understanding the content do not hinder its qualification as a translation.* (EPO, 2018, Part G, Chap. IV-4)

Patent examiners typically share the results of their patentability search with patent applicants, and any relevant patent document that is in another language is provided in machine-translated form. Unless the applicant decides it is so important that they will provide a human

[1] Here and elsewhere: PP = Patent Professional. Also, quotes have been edited for fluency.

[2] Some quotes and passages translated by author.

translation, prosecution proceeds based on the machine translation. MT is occasionally used in opposition proceedings as well:

PP9*: I mean normally in opposition cases at the EPO, European Patent Office, you can use machine translations.*
Interviewer*: OK. And in the Finnish patent office as well?*
PP9*: Yes you can do that. I have never been asked to provide a human translation about any of these.*

5 Deciding on human translation: an exercise in risk assessment

As far as translation is concerned, the most important decision patent professionals or patent applicants make is whether to rely on gist MT for understanding a relevant document or to have it translated by a human. Nitzke et al. (2019) proposes that this type of decision involves risk and that an assessment of those risks should be part of the process of decision-making. Evidence of such a risk assessment emerged in this study, with patent professionals weighing various factors before deciding on gist MT or human translation. The factors that supported human translation of a patent document included the riskiness of the IPR process in which the document would be used, the assumed relevance and importance of the other-language document, and the potential consequences if a misunderstanding would occur due to an error in the gist MT. The factors supporting the use of gist MT were lower costs, quicker access to information, and trust that the patent document is adequately understood. This assessment was summarized by some informants:

…the more important decision, the less you do the decision based only on the machine translation. (PP8)

…if the context is clear then it's OK as I see it, I trust the machine translations enough, but sometimes when we are in borderline decisions it's required to have a proper human translation…So it's more a question of the uncertainty margin of the translation with respect to what we are deciding. (PP4)

Each side of this assessment is examined more closely below.

5.1 Arguments for human translation

One of the top considerations for triggering human translation was the IPR process the other-language relevant document would be used in, with some processes being seen as more high-risk than others. Cases that involved infringement or freedom to operate might involve considerable costs and legal involvement, and these were consequently cited as cases in which human translation is often needed:

It depends on the costs involved in the case…if you are in a patent battle, if there is an infringement case…there's a lot of money involved. If you want to be absolutely sure then you have to have a human translator. (PP9)

If, based on the gist MT, a patent document appears to be highly relevant and important to a case, that would also serve as a strong argument for triggering human translation:

…probably that also depends a little bit on the case. If it is highly important then I would choose immediately to get it translated, or claims or parts of it, translated with human translation. (PP2)

Informants also mentioned potential consequences as a factor in the decision on human translation:

If we make the wrong decision and allow a product to the market which does not have freedom to operate, there is a risk of using time and money and goodwill in a court case and potentially being responsible to cover the damages of a client. (PP4)

5.2 Arguments for relying on gist MT

The main arguments for using gist MT are clearly that translation is very quick and does not generate extra costs. MT is provided at no cost by various national and international patent offices such as the Japan Patent Office and the EPO, and it is commonly included by default in commercial patent search tools. Its use is also made easy through tight integration to patent search tools and processes.

A complete understanding of the arguments for relying on MT in the risk assessment, however, requires consideration of another important element: how strong is the patent professional's trust that they have a sufficient understanding of a patent document? Much of this depends on the quality of the machine translation, of course. However, past studies have shown that other factors can enhance users' abilities to understand MT, and those were reported as helpful in this study as well. Two factors appeared to contribute to trust in understanding in this case study: the fact that patent professionals rely on other resources than the gist

MT, and the background knowledge that patent professionals possess. These are discussed further in the following subsections.

Understanding does not depend on MT alone

The understanding of the machine translation of a patent document can be seen as a process of trying different alternatives until a sufficient level of understanding is achieved. The first alternative is to combine the gist MT with other resources, such as drawings and chemical formulas in the original-language patent documents, to enhance understanding. This combining of MT and auxiliary, often multimodal, information to obtain an understanding of other-language texts has been reported in other studies on MT users (Nitzke et al., 2019; Pituxcoosuvarn et al., 2018; Suzuki and Hishiyama, 2016; Way, 2013; Gaspari, 2004; Henisz-Dostert, 1979). Auxiliary information had a clear role in patent professionals' reports of their work in this study as well:

> *When it's good enough that I can see that it's relevant? It's a combination of understanding the figures and understanding the machine translated text.* (PP6)

> *Oh yeah then you have to look at the original because it doesn't translate any of those chemical formulas…And then if they are totally different then it could be that I don't even make any translation because then I know that, well, they are talking about totally different things.* (PP7)

A second alternative patent professionals resort to are alternative machine translations from other MT tools, a practice that has been noted in earlier studies (Gao et al., 2015; Tinsley et al., 2012). At least one commercial patent search tool offers users access to both their own MT solution and the alternative of Google Translate in the same window. Although a few informants mentioned using a general tool such as Google Translate for alternative translations, a more common method was to try the MT tools provided by specific governmental patent offices:

> *I do the EPO machine translation first and if that's not more understandable then I go directly to the patent office that the publication came from, so Chinese or Japanese.* (PP5)

> *…for instance if it's a Chinese document I go to Chinese Patent Office website and try to find the same application there…and usually it's a different machine translation and that actually helps sometimes, when you have two machine translations you can read them at the same time and maybe it gives you a better impression.* (PP6)

The next alternative professionals can turn to are the other patent professionals they collaborate with. Instead of ordering a human translation of a text that is not sufficiently understood, they can ask the patent professionals who work more closely with the inventors (for example, the patent professionals in the country which the patent originates from) to clarify unclear passages for them.

Background knowledge aids understanding

As mentioned previously, the informants of this study were both highly experienced in the IPR field and well educated. Their contextual knowledge and competences in languages appeared to be important factors in helping them understand and use machine-translated information effectively.

The importance of MT users' knowledge of context in helping them understand machine-translated texts has been reported in a number of studies. Henisz-Dostert (1979) found that a user's familiarity with the subject matter was seen as the main factor in determining the understandability of machine-translated texts. Other studies that have highlighted the importance of contextual knowledge include Bowker and Buitrago (2019), Yasouka and Björn (2011), Yamashita et al. (2009) and Smith (2003).

In the patent context, contextual knowledge is often divided between the patent professionals, who know the patent genre, and the inventors or researchers behind the patents, who know the subject matter better. These competences, their role in helping to understand machine-translated patents, and the division of expertise between patent professionals and inventors were a common theme in the interviews.

> *And when you understand…if we're talking about patent publications there's a certain structure and there's a certain format that they're in, then it's in a way easy easier to follow.* (PP2)

Several previous studies have examined the role of users' language competence in gist MT scenarios (Nurminen and Papula, 2018; Nurminen, 2016; Henisz-Dostert, 1979). In the current study, this background competence also appeared to be a factor in successful use of MT. Although none of the informants spoke English as their native language, all used English daily in their work. Their MT use was mainly from other

languages into English, not into their native languages. Besides English, all informants had varying levels of competence in other languages, with German being the most often mentioned, followed by French, Spanish, Swedish, Italian, Dutch, and Japanese. Several informants indicated that competence in the source language helped them to understand texts that were machine-translated from those languages:

> *And quite often I actually combine a machine translation and original reading…the complementarity of understanding the structure of the language better than the machine, and the machine understanding more words than I do, is a good complementarity.* (PP4)

However, the reality is that the major languages patent documents are translated from are Chinese, Japanese, and Korean because these countries are significant producers of patents. China became the world's largest patent producer in 2011. By 2017, China had filed 1.3 million patent applications, more than double the number filed by the second country, the U.S. (WIPO, 2018: 40). The predominance of China was mentioned in all interviews. We can assume from this that competence in the three major patent languages of Chinese, Japanese, and Korean would be particularly useful for patent professionals.

6 Affordances in the patent context

Thus far this article has presented a scenario in which patent professionals can and do use gist MT to make certain decisions. The article has also discussed the factors involved in their decisions to rely on gist MT or to order human translation. However, in the analysis of this study's data, certain contextual factors emerged which appeared to make affordances for the use of gist MT. These affordances must be considered when discussing this specific case because they appear to play an important role in making the use of gist MT tenable, and an understanding of this ecosystem's use of gist MT is incomplete without them. The following sections explore two factors of affordance, risk tolerance and legitimacy.

6.1 Risk-tolerant environment

In the book titled *Translation Quality Assessment*, Andy Way states that MT systems need to be evaluated with the knowledge of what the system would be used for. Way also notes that "[o]f course, some objectives could be more tolerant of MT errors than others" (Way, 2018: 170). Certain features of the patent environment, while perhaps not fully error-tolerant, appeared to make affordances for the risks and potential errors tied to the use of gist MT.

The patenting process is long and iterative, with multiple parties often reviewing the same or similar texts. Different stakeholders may have different interpretations of a patent application's scope and claims. To address these issues, the process contains space for discussion and mechanisms for stakeholders to examine and challenge each other's work. One of these is explained in the Finnish Patent and Registration's Patent Guide:

> *Even though inventions must show absolute novelty, it is not possible for patent authorities to clarify all public information when examining an application. For this reason, the examination process is augmented by the third-party observation and opposition processes, in which third parties, for example competitors, can bring to the attention of the authorities issues that did not emerge during the examination of the patent application. (PRH, 2018: 19)*

The nature of this process means that there are also multiple stages where errors in the understanding of gist MT can be detected and corrected. This was described by one informant:

> *Well of course you can get the wrong impression of the subject matter in the document, but I don't see that it's a really big risk because the patent application process is a long process, so if my interpretation of some kind of document based on the machine translation is wrong, I can change my mind later, if I see it. It takes usually over two years to get a patent so we get the answer from the applicant and we probably write another office action and then the applicant replies again, so it's a conversation. So during the process there's many times when these things can be dealt with.* (PP6)

Another informant described a case when parties examined and challenged each other's MT work:

> *We've had these cases where the examiner used Google Translate and we translate it using EPO's official site and then we can explain to them that 'now we would like to kindly point out that the translation used by the examiner contained a mistake in this spot, and that we have this in that same spot, and our version uses the terms in this way.' And we rely on the [machine] translation completely…the examiner doesn't understand Japanese and we don't understand Japanese. We are both relying on machine translation and there is nothing else.* (PP1)

Besides the risk tolerance present in processes, the informants in this study displayed a tendency to accept the risk involved with using MT and making decisions based on it. One reason for this might be that the informants were vastly experienced. Another reason might be that the IPR field contains other risks besides the use of gist MT, so the organizations they work in might have a higher willingness to take risks, or "risk appetite" as defined by Nitzke et al. (2019). Finally, the acceptance of risk might be an acknowledgement that the risk is simply necessary due to the impossibility of relying on human translation for the large volumes of documentation they regularly encounter, as voiced by one informant:

> *Yes, there is always risk involved. But we have so many patents to go through. Hundreds and hundreds at a time. It would be impossible if all of those had to be translated by a human. Always a risk though.* (PP3)

6.2 Legitimacy of MT

One aspect of the use of MT in the patent environment that I did not expect when I began my research was the legitimacy that it enjoys. One of Merriam-Webster's definitions for *legitimate* best reflects what it means in this context: "conforming to *recognized* principles or *accepted* [emphasis by author] rules and standards."[3] Three different themes in this study illustrated this legitimacy: MT use was transparent, the boundaries of its legitimacy were documented and generally agreed upon by users, and its users had a relatively high level of 'MT literacy.'

Transparency

Transparency in gist MT use has been addressed in a few reports, most recently in a 2019 Globally Speaking Radio podcast in which John Tinsley reported that the legal profession is beginning to use MT for e-discovery, and that its use is fully transparent in that context: "So you go into the court and say to the judge, 'We are taking this position on the basis of a machine translation of this document into English,' and that's legally defensible" (Beninatto and Stevens, 2019).

At least in the European context of this study, the first sign of transparency was the inclusion of MT in EPO guidelines. Second, descriptions by study informants depicted an environment in which the role of MT is transparent to all. They also reported that MT is transparent to secondary users of patent MT, the internal and external clients the patent professionals work with. The results of searches these clients receive from patent professionals often include documents that are machine-translated. These are clearly marked as machine translations and they often also include the date and MT tool that produced the text. Patent professionals discuss MT with the secondary users, as in this example:

> *I would point out that this is a machine translation and, depending on if the client is a knowing patent engineer, then I would maybe give my opinion if we need a proper translation or not, but then ask what they think.* (PP8)

Boundaries of legitimacy

An important aspect of legitimacy is that it is bound to a specific scope. The 'recognized principles or accepted rules and standards' referred to in the definition provided earlier are agreed upon by a certain group for a certain purpose, and the boundaries of applicability are recognized by the participants. In this study, the boundaries of legitimacy for MT were sometimes mentioned during answers to other questions: "For information purposes, it's fine. For use as a legal text, no." (PP3) But in the interviews I also asked directly, "In what situations is it not OK to use machine translation?" The responses indicated some agreement on the areas in which MT should not be used, such as when filing patent applications:

> *...when you're translating your application to other languages – like we have seen some kind of, I think they are usually applicants from Asia, that file an application here and usually they apply in English, but you can really see that their application is machine translated from the Chinese version – not OK.* (PP6)

There was very clear agreement that MT should not be used in legal settings, as in this example in which an informant described a process involving another company's potential infringement of their patent:

> *We would start with searching prior art and use MT. Our aim is to see if there's overlap with our patent or not. If we find something that looks in-*

[3] https://www.merriam-webster.com/dictionary/legitimate

teresting, then we would order human translation of it. We would not just go ahead on that with machine translated information. (PP3)

MT literacy

In 1993 Church and Hovy defined six "desiderata" for a good use case for MT. Among the six were: "it should set reasonable expectations" and "it should be clear to the users what the system can and cannot do" (Church and Hovy 1993: 257). Bowker and Buitrago (2019) expanded this idea by coining the term *MT literacy,* and then applying it to the case of MT use in academic work. On the basis of their definition, I described MT literacy for the context of this study as a patent professional's ability to: (1) comprehend the basics of how machine translation systems process texts, (2) understand how machine translation systems are or can be used (by oneself or others working with patents) to find and read patent documents within the context of IPR processes, and (3) appreciate the wider implications associated with the use of machine translation. Based on this definition, the informants in this study displayed a generally high level of MT literacy. They appeared to understand the basics of MT technologies, knew how to access different MT tools, and were aware of the possibility and consequences of translation errors. They also had experience with different types of MT tools and noticed improvements in quality over time:

They all [languages] *have become better, and especially nowadays if you make a machine translation for some 'normal' language, for German or French, they are really good.* (PP9)

Perhaps one of the clearest signs of the high level of general MT literacy was an observation I made throughout the study: the hype issues currently visible in other spheres (for one example, see Hassan et al., 2017 followed by Toral et al., 2018) do not seem to be occurring in patent MT. In the present study, MT was considered to be one tool among others and people were aware of its uses and limitations. I heard no reports of overreaching claims on MT capabilities.

7 Conclusions

The main objective of this study was to explore the types of decisions patent professionals make based on machine-translated information, the risk assessment they use when deciding between relying on gist MT or using human translation, and the environmental factors that appear to support the use of gist MT in this context. The results revealed that patent professionals routinely make decisions on relevance and monitoring based on gist MT, and that the patenting process also relies on it. In the key decision of initiating human translation, patent professionals tend to weigh the riskiness of the IPR process in which the translated patent document would be used, the assumed relevance and importance of the document, and the potential consequences of misunderstanding against the lower costs, quicker access to information, and trust in a good enough understanding of the patent document. That understanding is often based not only the gist MT, but also other factors, such as auxiliary information sources and patent professionals' contextual and linguistic knowledge. The environmental factors of risk tolerance and legitimacy for gist MT also support the use of MT.

The study contributes to our knowledge of how people, and specifically professional groups, use gist MT. It explores factors that can enhance the use of gist MT, and this understanding will help us to define the characteristics of good, as well as poor, contexts for gist MT use. In addition, this analysis contributes to the growing body of research on users of various types of artificial intelligence.

This study had certain limitations. The group of informants was small and somewhat homogeneous, and this influenced the results. Data was gathered through only one method, interviewing. The results also focused on patent work in one geographical area and one specific point in time, and the results cannot be considered representative of the larger population of patent professionals. Nevertheless, as the first exploratory study of this very experienced group of MT users, it fulfilled one of the main purposes of inductive research in that it uncovered new themes and hypotheses on how a specific group uses gist MT and on the contextual factors that contribute to their use of it.

Further studies on this gist MT user group could target an expanded group of informants, including more diverse participants, other patent MT user groups, and less experienced patent MT users. Studies incorporating other methods such as contextual inquiry, diaries, or quantitative methods such as surveys could verify some of the findings of this study and might reveal further insights on this user group. In addition, it is hoped that we will see a growth in the research, and number of researchers, devoted to studying all types of users of gist MT.

References

Alberts, Doreen, Cynthia B. Yang, Ken Koubek, Suzanne Robins, Matthew Rodgers, Edlyn Simmons, and Dominic DeMarco. 2017. Introduction to Patent Searching. In Lupu, Mihai, Katja Mayer, Noriko Kando and Anthony J. Trippe, editors, *Current Challenges in Patent Information Retrieval*, Second ed. Vol. 29, Springer, Berlin Heidelberg.

Beninatto, Renato and Michael Stevens. 2019. What's the Latest with Neural MT? *Globally Speaking Radio* (podcast), episode 75. RWS Moravia.

Bowker, Lynne and Jairo Buitrago Ciro. 2019. *Machine Translation and Global Research: Towards Improved Machine Translation Literacy in the Scholarly Community,* Emerald Publishing, Bingley, UK.

Bowker, Lynne and Melissa Ehgoetz. 2007. Exploring User Acceptance of Machine Translation Output: A Recipient Evaluation. In Kenny, Dorothy and Kyongjoo Ryou, editors, *Across Boundaries: International Perspectives on Translation Studies*, Cambridge Scholars Publishing, Newcastle, UK.

Braun, Virginia and Victoria Clarke. 2013. *Successful Qualitative Research: A Practical Guide for Beginners.* SAGE, Los Angeles, California, USA.

———n.d. Thematic analysis: reflexive approach, accessed March-June, 2019. https://www.psych.auckland.ac.nz/en/about/our-research/research-groups/thematic-analysis.html

———2006. Using thematic analysis in psychology. *Qualitative Research in Psychology,* 3(2):77–101.

Burgett, Will. 2015. Unmoderated Remote Usability Testing of Machine Translation Content. *TAUS Review of Language Business and Technology,* IV:30–37.

Canfora, Carmen and Angelika Ottmann. 2018. Of ostriches, pyramids, and Swiss cheese: Risks in safety-critical translations. *Translation Spaces,* 7(2):167–201.

———2016. Who's afraid of translation risks? Presentation at 8th EST Congress, Aarhus, Denmark.

Church, Kenneth W. and Eduard H. Hovy. 1993. Good Applications for Crummy Machine Translation. *Machine Translation,* 8:239–258.

Dillinger, Mike and Laurie Gerber. 2009. Success with Machine Translation: Automating Knowledge-Base Translation, Part 1. *ClientSide News*, 10.

EPO: European Patent Office. 2018. Guidelines for Examination in the European Patent Office. European Patent Office, Munich, Germany.

Gao, Ge, Bin Xu, David Hau, Zheng Yao, Dan Cosley, and Susan R. Fussell. 2015. Two is Better than One: Improving Multilingual Collaboration by Giving Two Machine Translation Outputs. *18th ACM Conference on Computer Supported Cooperative Work & Social Computing,* Vancouver, BC, Canada, 852–863.

Gaspari, Federico. 2004. Online MT Services and Real Users' Needs: An Empirical Usability Evaluation. *6th Confererence of the Association for Machine Translation in the Americas, AMTA 2014,* Washington DC, USA 74–85.

Hassan, Hany, Anthony Aue, Chang Chen, Vishal Chowdhary, Jonathan Clark, Christian Federmann, Xuedong Huang, Marcin Junczys-Dowmunt, Will Lewis, Mu Li, Shujie Liu, Tie-Yan Liu, Renqian Luo, Arul Menezes, Tao Qin, Frank Seide, Xu Tan, Fei Tian, Lijun Wu, Shuangzhi Wu, Yingce Xia, Dongdong Zhang, Zhirui Zhang, and Ming Zhou. 2018. Achieving Human Parity on Automatic Chinese to English News Translation. https://arxiv.org/abs/1803.05567

Henisz-Dostert, Bozena. 1979. Users' evaluation of machine translation. In Winter, Werner, editor, *Machine Translation*. Mouton Publishers, The Hague.

ISO. 2010. Ergonomics of Human-System Interaction. Part 210, Human-Centred Design for Interactive Systems (ISO 9241-210:2010*)*. Suomen standardisoimisliitto SFS, Helsinki, Finland.

Joho, Hideo, Leif Azzopardi, and Wim Vandervauwheded. 2010. A Survey of Patent Users: An Analysis of Tasks, Behavior, Search Functionality and System Requirements. *3rd Information Interaction in Context Symposium*, New Brunswick, NJ, USA 13–24.

Larroyed, Aline A. 2018. Machine Translation and Disclosure of Patent Information. *IIC - International Review of Intellectual Property and Competition Law,* 49(7):763–786.

List, Jane. 2012. Review of machine translation in patents - Implications for search. *World Patent Information,* 34(3):193–195.

Liu, Nancy Xiuzhi and Matthew Watts. 2019. Mobile Translation Experience: Current State and Future Directions. In Xu, Xiaoge, editor, *Impacts of Mobile use and Experience on Contemporary Society,* IGI Global, Hershey, PA, USA.

McDonald-Maier, Lisa. 2009. esp@cenet: Survey reveals new information about users. *World Patent Information,* 31(2):142–143.

Merriam, Sharan B. and Elizabeth J. Tisdell. 2016. *Qualitative Research: A Guide to Design and Implementation*. John Wiley & Sons, San Francisco, CA, USA.

Nitzke, Jean, Silvia Hansen-Schirra, and Carmen Canfora. 2019. Risk management and post-editing competence. *The Journal of Specialised Translation,* (31):239–259.

Nurminen, Mary. 2016. Machine Translation-Mediated Interviewing in Qualitative Research: a Pilot Project. *New Horizons in Translation Research and Education*, 4:66–84.

Nurminen, Mary and Niko Papula. 2018. Gist MT Users: A Snapshot of the Use and Users of One Online MT Tool. *21st Annual Conference of the European Association for Machine Translation*, Alicante, Spain 199–208.

Oesch, Rainer, Heli Pihlajamaa, and Sami Sunila. 2014. *Patenttioikeus* [Patent Law]. 3rd ed. Talentum, Helsinki, Finland.

Pituxcoosuuvarn, Mondheera, Toru Ishida, Naomi Yamashita, Toshiyuki Takasaki, and Yumiko Mori. 2018. Machine Translation Usage in a Children's Workshop. *10th International Conference on Collaboration Technologies,* Costa de Caparica, Portugal, 57–73.

PRH: Patentti- ja rekisterihallitus [Finnish Patent and Registration Office]. 2018. *Patenttiopas* [Patent Guide].Patentti- ja rekisterihallitus, Helsinki, Finland.

Pym, Anthony. 2015. Translating as risk management. *Journal of Pragmatics,* 85(Aug):67–80.

Pym, Anthony and Kayo Matsushita. 2018. Risk Mitigation in Translator Decisions. *Across Languages and Cultures,* 19(1):1–18.

Rossi, Laura and Dion Wiggins. 2013. Applicability and application of machine translation quality metrics in the patent field. *World Patent Information,* 35(2):115–125.

Smith, Ross. 2003. Overview of PwC/Sytranet on-line MT Facility. *Twenty-Fifth International Conference on Translating and the Computer*, London, UK.

Stewart, Osamuyimen, David Lubensky, Scott Macdonald, and Julie Marcotte. 2010. Using Machine Translation for the Localization of Electronic Support Content: Evaluating End-User Satisfaction. *The 9th Conference of the Association for Machine Translation in the Americas,* Denver, Colorado, USA.

Suzuki, Hiroshi and Reiko Hishiyama. 2016. An Analysis of Expert Knowledge Transmission using Machine Translation Services. *Seventh Symposium on Information and Communication Technology*, 352–359.

Tinsley, John. 2017. Machine Translation and the Challenge of Patents. In Lupu, Mihai, Katja Mayer, Noriko Kando and Anthony J. Trippe, editors, *Current Challenges in Patent Information Retrieval*, Second ed. Vol. 29, Springer, Berlin Heidelberg.

Tinsley, John, Alexandru Ceausu, Jian Zhang, Heidi Depraetere, and Joeri Van de Walle. 2012. IP-Translator: Facilitating Patent Search with Machine Translation. *AMTA-2012: Tenth Biennial Conference of the Association for Machine Translation in the Americas*, San Diego, CA, USA.

Toral, Antonio, Sheila Castilho, Ke Hu, and Andy Way. 2018. Attaining the Unattainable? Reassessing Claims of Human Parity in Neural Machine Translation. *Third Conference on Machine Translation*, Brussels, Belgium 113–123.

Way, Andy. 2018. Quality Expectations of Machine Translation. In Moorkens, Joss, Sheila Castilho, Federico Gaspari & Stephen Doherty, editors, *Translation Quality Assessment: From Principles to Practice*, Springer, Cham, Switzerland.

Way, Andy. 2013. Traditional and Emerging use-Cases for Machine Translation. *Translating and the Computer,* London, UK.

WIPO: World Intellectual Property Organization. 2018. World Intellectual Property Indicators 2018. World Intellectual Property Organization, Geneva, Switzerland.

Yamashita, Naomi, Rieko Inaba, Hideaki Kuzuoka, and Toru Ishida. 2009. Difficulties in Establishing Common Ground in Multiparty Groups using Machine Translation. *CHI 2009, the SIGCHI Conference on Human Factors in Computing Systems*, Boston, MA, USA 679–688.

Yasouka, Mika and Pernille Bjorn. 2011. Machine Translation Effects on Communication: What Makes it Difficult to Communicate through Machine Translation? *IEEE*, Kyoto, Japan 110-115.

PUBLICATION III

Raw machine translation use by patent professionals: A case of distributed cognition

Mary Nurminen

Translation, Cognition and Behavior, Volume 3(1): 100–121
https://www.benjamins.com/catalog/tcb

Nurminen, Mary. (2020) Raw machine translation use by patent professionals: A case of distributed cognition. *Translation, Cognition and Behavior*. 3(1): 100–121.
doi: https://doi.org/10.1075/tcb.00036.nur

Raw machine translation use by patent professionals

A case of distributed cognition

Mary Nurminen
Tampere University

This article examines the use of raw, unedited machine-translated texts by patent professionals using the framework of distributed cognition. The goals of the study were to evaluate whether the concept of distributed cognition is a useful theoretical lens for examining and explaining raw MT reception, and to contribute to our knowledge of raw MT use through an analysis of a real-life use case. The study revealed that patent professionals often rely on a large network of artifacts and people to help them in the task of understanding raw MT, and therefore the concept of distributed cognition was applicable and useful. The study also contributed new knowledge to our overall understanding of the use of raw MT.

Keywords: raw machine translation; machine translation gisting; patent machine translation; machine translation users; distributed cognition; machine translation for assimilation

1. Introduction

The use of raw, unedited machine translation (MT) to acquire a basic understanding of a text is today a widespread phenomenon, with the number of users of Google Translate alone estimated in 2016 to be 500 million (Turovsky 2016). Raw MT is also being used in professional contexts, not only in professional translation but also in fields such as customer support and academia. As reported by Nurminen (2019), raw MT has been used on a broad scale in the field of intellectual property rights (IPR) for approximately a decade. Patent professionals, including patent counsels, attorneys, examiners and other professionals, use raw MT to read patent documents that are in languages they know only a little or not at all.

The case of raw MT use by patent professionals has gone largely unnoticed by researchers. In fact, our knowledge of the reception of raw MT in general is still lacking, and very few hypotheses or theories have been developed on the cognitive aspects of raw MT reception. There is a considerable body of research on cognition and translation (e.g., Shreve & Angelone 2010; Schwieter & Ferreira 2017), much of which might be applicable to the context of raw MT use, but few studies explore the application of that knowledge to raw MT cognition.

This article analyzes one case of MT gisting—the use of raw MT by patent professionals (definition in next paragraph)—through the lens of the theoretical framework of distributed cognition. My first goal in this article is to apply the principles of distributed cognition to the case study, evaluating whether it is a useful means for explaining raw MT reception. My second goal is to build on our understanding of the nature of MT gisting through this analysis of a real-life use case, so that we can begin to understand some of the conditions that are present in a functioning use case of MT gisting. In this inductive, exploratory study, the intent is not, and cannot be, to establish causality. Rather, the study aims at an initial identification of factors and conditions that appear to influence the use of raw MT, either positively or negatively. Those factors can then be further explored and tested in future research.

A few notes on terminology are in order. The first concerns languages. *L1* refers to the language a person has the best command of, which is often the language they use most often in their everyday lives. *L2* indicates a language a person has a very good command of and uses on a

regular basis. *L3* applies to other languages, which might be languages a person has some competence in or languages they do not understand at all. The Scandinavian informants in this article represented three different L1s and all of them have English as an L2. They use it daily in their work and the raw MT texts they work with are translated from other languages into English.

The second note concerns the phenomenon in which a person reads raw, unedited MT output to gain at least a basic level of understanding of it. In this article, that phenomenon is referred to as *MT gisting*, while the text that is being read in MT gisting is termed either *raw* or *unedited* MT and the person reading raw MT for gisting purposes is referred to as a *raw MT user*. Finally, in this study, the term *patent professionals* refers to professionals who use their expertise in IPR processes to assist internal or external clients with their Intellectual Property Rights needs.

The next section reviews literature related to this study. Section 3 describes the methods used in the case study on patent professionals. Readers are not necessarily familiar with IPR work, so Section 4 introduces the environment in which patent professionals work. Section 5 discusses the results of the analysis, and conclusions are offered in Section 6.

2. Related work

Very little research has been done in the cognitive specifics of MT gisting. O'Brien (2017) devotes a chapter to MT and cognition in *The Handbook of Translation and Cognition*, but she comments frankly in the beginning that, since research on users of raw MT is in its early stages, it would receive little focus and the section would be devoted mostly to post-editing and evaluation processes. The one mention of raw MT users is in relation to evaluation methods, and the paragraph warns that the methods covered in the section:

> […] are taken, to some extent at least, as acting as a proxy for measuring the usability and acceptability of MT output by those for whom it is essentially produced: end users. However, these are not direct measures of interaction with MT and tell us very little about the cognitive processing effort that might be involved in reading, comprehending, and acting on a set of machine-translated task instructions.
> (O'Brien 2017, 327–328)

However, a few studies on cognition and reception of various types of texts have included raw MT as one text type. O'Brien (2010) tested the effect of controlled language on readability for readers of texts in the original source language as well as in raw MT form in different languages. Other studies used eye tracking combined with other methods, such as post-task questionnaires, to investigate end users' evaluations of factors such as usability and acceptability in texts (Castilho et al 2014; Doherty & O'Brien 2009, 2014; Castilho & O'Brien 2017). These studies again tested different types of texts, including original texts, human translations, post-edited MT and unedited MT, and much of the focus of the studies was on comparing the results of those different types of texts, rather than on specifically exploring the cognitive aspects of raw MT reception.

2.1. *Distributed cognition and translation*

An important area of background research I relied on when writing this article dealt with situated and distributed cognition in the processes of translators. One might assume that it would make sense to use studies on areas closer to raw MT reception; for example, studies on cognition and reception of human translation. And, as O'Brien (2017) noted in the quote above, the cognitive situations of a translator or post-editor at work and a user of raw MT are different. However, it is in the area of situation cognition and translation that the greatest theoretical development has taken place and for that reason, material from that area was found to be the most helpful.

Muñoz (2010, 172) states that "cognition takes place in the context of inputs and outputs relevant to the task at hand." He applies this to the work of translators and the way they "act together to accomplish the complex task of creating a single, communicative product" (ibid.) Risku (2014, 335) discusses "situated, embodied and extended cognition" and how "cognitive processes are context-dependent, i.e., they are dependent on and partly constituted by the social and physical environment in which they are carried out." Muñoz (2017, 564) also illustrates how situation can affect meaning: "Meaning is encyclopedic, and it is a process, not a thing […] Understanding is an activity that crucially depends on the environment—and also on experience—because environmental affordances foster and constrain meaning construal." Although all of these descriptions are in articles about translators' cognition while translating, they are fitting descriptions of the work in this use case of MT gisting.

The work in situated cognition in translation studies has also led to a new emphasis on qualitative research methods. Risku (2014, 335) argues that, while the controlled experiment research that is dominant in translation process research is necessary, "we also need other methodological paths of inquiry to model the cognitive processes in translation and to establish a deeper understanding of how translations are produced." Ehrensberger-Dow (2019, 41) also emphasizes that study in this area relies on both experiments in laboratory settings and workplace observations.

2.2. *Raw MT users*

The number of studies published on raw MT users remains relatively small. One area of research consists of surveys and market studies, some of which focus on the use and users of a specific tool (Smith 2003; Yang & Lange 2003; Nuutila 2005; Burgett 2015; Nurminen & Papula 2018). A few analyze specific types of users, such as Gaspari's (2007) survey of students and Henisz-Dostert's (1979) survey of scientists. The latter of these differs from the others in that, although it was a survey, it was a qualitative one, conducted as a series of structured interviews. It also included a few questions on cognitive processes, such as mental corrections of awkward translations.

Another group of studies involve laboratory experiments in which volunteer participants, who are often not actual users of raw MT, are asked to evaluate specific aspects of raw MT use. Some of these were mentioned in the previous section. Others include experiments on aspects such as comprehensibility, confidence, and acceptability (Fuji et al 2001; Gaspari 2006; Bowker & Eghoetz 2007, to name a few).

A third area of research addresses MT for communication, which refers to raw MT used to enable communication between people who do not share a language. Many of these studies are experimental, but they are carried out in extended virtual environments that simulate real-life environments. Several studies involve working groups based on different sites that have different L1s. The groups are assigned a task to accomplish together, with all communication surrounding the task occurring in chat applications backed by MT. The communication transcripts from the chats are then analyzed. Often these experiments test the influence of individual factors on communication. For example, Gao et al (2013) tested the highlighting of keywords. Xu et al (2014) and Gao et al (2015) examined scenarios in which users were presented with two different MT outputs. Yasouka & Björn (2011) and Yamashita et al (2009) examined strategies to establish common ground among the participants.

Only a few use cases of professional groups using MT gisting have been studied. Although scientists were the first group studied, the research by Henisz-Dostert (1979) remains the only one of its kind. Some scholars have examined the use of raw MT in customer support (e.g., Stewart et al 2010; Burgett 2015). There is a growing body of research on the use of MT by academics (for example, Bowker & Eghoetz 2007; Jolley & Maimone 2015; Bowker & Buitrago 2019). A few projects have also studied the use of raw MT in health care (review in Liu & Watts 2019).

A large number of the aforementioned studies on MT gisting are quantitative and/or experimental. Risku's (2014) call for more diverse methodologies in research could likewise be applied to raw MT reception, where our challenge is to model the cognitive processes so that we

can acquire a deeper understanding of how raw MT users understand raw MT and what strategies they adopt when faced with challenges in understanding.

3. Methods

This article is based on a case study project on patent professionals' use of raw MT. The aim of the project (Nurminen 2019) was to explore the context in which raw MT is used by the group. During the data analysis phase, it became evident that distributed cognition might be an effective theoretical framework for explaining the phenomenon and that realization led this article. The study was conducted by adopting inductive, qualitative methods. Especially in such an under-researched area, a more nuanced analysis was needed that would not only contribute to an overall understanding of the phenomenon but would also help us identify the factors that could be tested in future controlled experiments.

The data for the study was gathered by conducting a series of semi-structured interviews between April 2018 and February 2019, either at informants' workplaces or through Skype audio calls, which were easier to arrange than video calls. Interviewing was chosen primarily because the time commitment it required from the highly busy informants was only 1.5 hours on average. Several sources helped me prepare interview questions. The patenting guidelines from the Finnish Patent and Registration Office (PRH 2018) and the European Patent Office (EPO 2018), as well as general books on patenting such as Oesch (2014) and Alberts et al. (2017) were instrumental in helping me understand the practicalities of patent professionals' work. Joho et al's (2010) survey of patent professionals was also useful. I considered gathering further data through other qualitative methods such as diaries, but was assured by patent professionals that people in their profession would not be willing to devote the time required for such activities.

The informants of the study were nine Scandinavian patent professionals from three types of organizations: five worked in companies that actively file and prosecute patents; two were from Intellectual Property Rights (IPR) service providers; and two worked for a governmental patent office. The patent-professional informants of the study were well educated: all nine had Master's degrees and four also had PhD degrees. Their experience in the IPR field ranged from 9 to 30 years, and their reported estimates on MT experience ranged from 4 to 15 years.

A total of 12 hours of interviews were recorded, transcribed, and analyzed using the thematic analysis method defined by Braun & Clarke (2006; 2013). This included the use of a member check after the initial analysis. Preliminary findings were compiled and then reviewed by three volunteer informants to validate that the findings accurately reflected their experience. Comments and clarifications were then considered in producing the final analysis. For this article, the themes and data were analyzed through the lens of distributed cognition.

4. Introduction to the work of patent professionals

This section briefly describes the aspects of the work conducted by patent professionals that relate to their use of raw MT, beginning with the cognitive activity under investigation. In analyses of translation and cognition, the activity that translators are involved in is the act of translating. In studies on reception and cognition, the activity readers are engaged in is reading a human translation. The activity under investigation in the case of patent professionals and cognition is one of reading and trying to understand a raw MT text.

4.1. *Texts and processes*

When inventors work on an invention that they are considering patenting, they must ensure that it is actually a new idea that does not exist, either in granted patents or in patent applications. Consequently, a large part of the work of patent professionals involves searching for and reviewing existing patent documents of inventions closely related to the one in question, with the intent of ensuring nothing in its claims has already been claimed in another patent. The goal is to identify a small set of patent documents that are the most relevant to the case at hand, and

then to shape their own patent's claims so as to prevent infringement on any of them. The decision to include or exclude something into that set of relevant patent documents is one of the primary decisions patent professionals make (Nurminen 2019, 34–35).

Patent professionals and inventors are responsible for reviewing all existing patent documentation on inventions that closely resemble their own, including documents filed in other languages and in other countries. In 2012, an estimated 30% of a typical patent search resulted in patent documents in languages other than English (Tinsley et al 2012, 1); it is reasonable to assume that the percentage is now higher. As they may be searching for and reviewing a significant number of documents, sometimes "hundreds and hundreds at a time" (PP3), it would be impossible for patent professionals to acquire human translations of all of them, and they therefore rely heavily on MT in their review. [1]

However, patent professionals and inventors do have the option of turning to human translation at any point if it becomes necessary. As described in Nurminen (2019), the decision to use human translation involves an assessment of risk and benefits. If the patent document that is in another language is needed for an especially risky or important IPR process, or if the patent professional cannot understand the MT, human translation might be introduced early in the process. On the other hand, if the IPR process is less risky or if the understanding they have of the raw MT is good enough, they will proceed with their case using the relevant documents in raw MT form. The alternative of using human translation is, however, always available. A similar access to human translation when needed has been reported in other studies of raw MT use, for example, Ishida (2016) and Pituxcoosuvarn et al (2018).

4.2. *MT for patents*

The patent genre is challenging for MT systems. The genre is highly formulaic—patent applications and patents have a standard structure and a writing style that is occasionally (more or less jokingly) referred to as *Patentese*' as described by Tinsley et al (2012, 69): "[it] typically comprises a mixture of highly specific technical terminology and legal jargon and is often written with the express purpose of obfuscating the intended meaning." The style also commonly involves difficult syntactic structures. Considering the evidence that the use of controlled language can improve the quality of MT output (cf. O'Brien 2010; Marzouk & Hansen-Schirra 2018), Rossi & Wiggins (2013, 117) first reviewed the eight main components of controlled language: sentences are short, grammatically complete, use the active form, have simple syntax, and express only one idea; spelling and punctuation are correct; and texts use a limited lexicon. They then analyzed the typical features of patent language against these components and reported that patent language was incompatible with all eight.

Despite these challenges, MT systems designed specifically to handle patent documents have been developed for more than 20 years, beginning with the first implementations in the 1990s in Japan (Cavalier 2001) and the eventual launch of a free MT service in Japan. This was followed by MT services offered by the State Intellectual Property Office of China in 2008 (Wang 2009), the World Intellectual Property Organization in 2011 (Pouliquen 2016), and the European Patent Office in 2012 (Battistelli 2012). Commercial patent databases started to include MT in their solutions in the early-to-middle 2000s.

Patent MT tools handle a large variety of languages. The most frequently translated languages are Chinese, Japanese and Korean, primarily because China, Japan and Korea produce a significant number of patents and patent applications every year. For example, the top four countries filing patent applications in 2017 were China, with approximately 1.3 million; the USA, with 600,000; Japan, with 300,000; and the Republic of Korea, with 200,000 (WIPO 2018, 40). All informants in this study mentioned the predominance of these languages and the effects of working with them.

[1] In this and other quotes from informants, spoken data has been edited so that it conforms to written language conventions. The identifier *PP* refers to patent professional.

4.3. *Tool environment*

One of the most important tools patent professionals use in their work are patent databases, which aggregate patent documents from a number of different national patent offices and offer them through one user interface. The databases provide patent professionals with a broad coverage of patents, advanced search tools, and other features that facilitate the complex task of searching. MT is one of those features. Sometimes the patent database manufacturer produces the MT in a bulk process, so that patent documents are already in machine-translated English when they are loaded into the database. Others provide tools to perform on-demand translations. MT is so commonly included in these systems that informant PP2 described the feature as "very routine, normal. You don't even ask for it separately". This ubiquity makes it possible to integrate MT directly into IPR processes, and the reports in this study illustrated a tight integration of MT into IPR tools and workflows.

An important element of the technical environment is that it provides functionality for saving machine-translated documents. Way (2013, 2018), Moorkens (2017), and Nitzke et al (2019) suggest that good use cases for raw MT most often involve, and should involve, highly perishable texts. However, the patent use case involves texts which are very long-lived. In addition to that, the machine translations of those texts can be long-lived, with translations being saved, stored, and shared among multiple people. Our informants described different ways they saved and shared machine translated texts. Some mentioned copy/pasting machine-translated excerpts of patent documents into Word documents or Excel sheets to share with technical teams. Two informants also mentioned that they also post-edited the excerpts a bit before sharing them. Informants also reported that they routinely save entire machine-translated documents in Word or PDF format to share with their teams or clients. Patent examiners store machine-translated documents in the systems that they use to communicate with patent applicants, allowing the applicants to access them directly from the system.

The technical environment also uses indicators to ensure that readers are aware that what they are reading is machine-translated. Informants described indicators such as colored frames, stamps labeled *[machine translation],* and documents with both original texts and translations shown in consecutive paragraphs or in side-by-side columns. These indicators often include the name of the system that produced the MT as well as a time stamp, which is either provided automatically by MT systems or added by patent professionals. These extra pieces of information can be important because the use of an inferior MT system or a translation considered to be dated may trigger the patent professional to perform a new translation. The informants of this study displayed an awareness of the advancement of MT technology and the improvements it brings, and mentioned that they sometimes re-translated documents because a newer translation might bring better results.

5. Distributed cognition: understanding through interaction with network

As mentioned, patent professionals are engaged in the cognitive activity of understanding, which in some cases might even be described as *deciphering*, a machine-translated text. In the middle of this process is one person reading and attempting to understand one machine-translated text. If the MT output is of low quality, a person unfamiliar with the use of raw MT might easily consider the task to be impossible. However, what if we envision the same activity as a socio-cognitive exercise in the environment in which it actually occurs, with the patent professional having access to "the roles and cognitive contributions of all co-workers," as occurs for translation (Muñoz 2010, 179)? In that light the task—and the chance for success in achieving the task—appear to be very different.

Clark & Chalmers (1998) present the idea of cognition coming through a *coupled system,* or a combination of a human mind and an external entity. They argue (p. 8) that this combination can be considered a cognitive system in its own right in that:

> All the components in the system play an active causal role, and they jointly govern behaviour in the same sort of way that cognition usually does. If we remove the external component the system's behavioural competence will drop, just as it would if we removed part of its brain.

At the core of the task performed by the patent professionals in our case study—understanding a MT of a patent document—is the coupled system of one patent professional, with the knowledge and competences they bring into the situation, and at least one external entity, an MT engine. Without the MT engine, the patent professional could make little or no sense of a patent document in a different language. Thus, as reasoned by Clark & Chalmers, the external MT engine is an "active" external component. If it is removed, the behavior of the human patent professional changes drastically.

In the case of patent professionals, the distributed nature of cognition encompasses far more than this core coupled system. Similarly to the translation project environment reported on in Risku (2014), patent professionals also work within a large network of artifacts and people, and they often achieve an understanding of patent documents in other languages through interacting with the elements of this network. In fact, the task of achieving understanding can be considered a process of trying different alternatives, using different elements of the network, until a sufficient level of understanding is achieved (Nurminen 2019, 37). The workings of such a network are perhaps best illustrated by Muñoz (2017, 564): "Cognition is also often *distributed*, in that several cognizing and not cognizing agents conjointly perform complex tasks, such as translating and producing large digital texts."

This section examines the distributed network of artifacts and humans that patent professionals access for assistance in understanding machine-translated patent documents. Table 1 introduces the elements and it is followed by a discussion of each.

Element	**Description**
Original source document	The patent document in its original language, complete with multimodal elements such as drawings
Inventors or technical experts	Closest human element
Alternative machine translations	Machine translations retrieved through a second MT engine
Larger network of stakeholders	Other people who have an interest in the IPR case in which the machine-translated patent document is being used

Table 1. Elements in patent professionals' network that contribute to understanding machine-translated documents

5.1. *Original source document*

The source-language patent document is one of the first elements in the network that patent professionals use to augment their understanding of the MT of a patent document. These documents frequently contain non-textual material, such as drawings or chemical formulas, that are not language dependent. They are "universally understood and should not be overlooked as important tools for reviewing [patent] documents in any language" (List 2012, 195) The informants of this study described how they used those parts of the original document as follows:

(1) PP9: *Sometimes you just have to use the figures of the application and then you have to compare and see what it might be.*
PP7: *It's a combination of the original, if I see the chemical formulas or whatever they are using, because those abbreviations, they are not translated, like in carbohydrate variations, they are not translating those […] and then I have the original and I have the translation. So then I combine them.*

Depending on the language of the original document, patent professionals also reported relying on their own language skills to use the source document. They all used English in their daily work and as a target language for MT, although none spoke English as their L1. In addition, all reported having some competence in German and most also had some command of other languages, such as French and Swedish. They reported reading patents in the original language when they felt their L3 command was good enough, but they also described situations in which they combined what they apprehended in the original with their grasp of the MT output to arrive at an understanding.

(2) PP2: *Well there's the thing that when you understand German, you check the original as well…a little bit. So that kind of helps you understand because you're not just solely reading the machine translation […] it of course helps because you can understand the structure quite well. You basically just need support with a few words you're unfamiliar with.*

Other studies have claimed a similar beneficial effect of language competence for raw MT reception. They have focused on the influence of source language command (Henisz-Dostert 1979; Morland 2002; Gaspari 2004; Nurminen 2016; Nurminen & Papula 2018), on target language skills (Henisz-Dostert 1979; Morland 2002; Smith 2003), or on proficiency in a pivot language (Ogura et al 2004).

The information sources listed thus far—the background knowledge of the patent professionals, the machine-translated text, and the original patent document—were very often enough to enable informants to acquire a satisfactory understanding of a document. One of the questions I asked in all interviews was, "Of all the times you use machine translation, how often is it successful, meaning you get enough information from the translation that you are immediately able to act on it?" A large majority reported that the success rate was very high. In six interviews, informants estimated their success rate at 90% or above.[2] One gave a somewhat lower rate, 75%:

(3) PP6: *It's a combination of understanding the figures and understanding the machine translated text, so I would say that maybe 75%? In those cases I'm pretty convinced if this document is relevant or not.*

Another gave an even more conservative estimate, stating that every other case required more checking. Despite these high success rates, the informants reported an occasional need to rely on other elements in the network to acquire understanding.

5.2. *Inventors and technical experts*

The humans in the patent professional's network that they work with most closely are the inventors or technical experts who have created the invention that requires a patent or is already protected by a patent. Patent professionals often work in an advisory role to the technical experts, assisting them through patenting and other IPR processes. In some scenarios, these two professional groups work in the same company, and the *client* of the patent professional is *internal*. Other patent professionals, such as those who work in patent service providers and governmental patent offices, work for external clients.

In some cases, the technical experts perform the initial task of searching through patents and selecting the most relevant ones; in others, it is the patent professional who does it. In either

[2] One of the informants reported that the initial 90% success rate had dropped after the technical field they worked in changed, resulting in an increased need to review documents from China and Japan.

case, both sides share the responsibility for understanding L3 patent documents, and this is often achieved through cooperative discussions. Sometimes, those are informal conversations between two colleagues, but informants also described reviews of machine-translated texts in meetings:

(4) PP1: *But many times we are in a meeting. I have the people there. I have [the tool] open and we are reading from it [...] and then when there's a German or French or Chinese text, we realized that oh, that's what we're dealing with, and the button is there and ready, Translate, and then it does it.*

When the patent professionals were working with external clients, these dialogues were described as more formal, but the process of achieving understanding through dialogue between various participants was similar. Even in the formal situation of the governmental patent office examining patent applications, understanding appeared to be an exercise in distributed cognition:

(5) PP6: *So if my interpretation of some kind of document based on the machine translation is wrong, I can change my mind later [...] It takes usually over two years to get a patent. So we get the answer from the applicant and we probably write another office action and then the applicant replies again. So it's a conversation.*

One factor that appeared to contribute to this patent professional-inventor pair's ability to understand MT documents was the combination of their different areas of contextual expertise, with patent professionals knowing the genre of patents and technical experts knowing about the subject matter or technical area of the patent in question. The informants in this study were experts in the patent genre due to their average of 17 years' experience in the IPR industry. They reported this to be helpful in understanding MT:

(6) PP2: *And when you understand the… if we're talking about patent publications, there's a certain structure and there's a certain format that they're in. Then it's in a way easier to follow.*

While they had the best understanding of the genre, the informants indicated that the technical experts were the "best expert[s] in the technical field" (PP6). Many also pointed out that it was the combination of their own expertise in the patent genre and the inventors' competence in the subject matter that was important:

(7) PP1: *So the system goes that way that, when we got the search results I send them out to the technical experts. They read them first by themselves. And they pick out those that they are worried about or where they want to have more information, where they're not sure what it really means. So then we have a meeting and we go through them together. So I can tell them what it really means and how to read it.*

Other studies have similarly noted the influence of contextual knowledge on the use of unedited MT. In Henisz-Dostert's (1979) study of scientists, informants were asked what the understandability of MT primarily depends on. They were given the choices of: (a) familiarity with subject matter; (b) translation of words; (c) sentence structure; (d) format; and (e) general style. The most cited option, by an overwhelming majority, was *familiarity with subject matter* (Henisz-Dostert 1979, 189). The influence of contextual knowledge on raw MT reception has also been discussed by Smith (2003), Yamashita et al (2009), Yasouka & Björn (2011) and Bowker & Buitrago (2019).

5.3. *Alternative machine translation*

The next element patent professionals use to gain understanding of machine-translated texts is output from an alternative MT engine. The patent professionals either compare two different outputs and select the better one, or they combine what they understood of both to construct an overall general understanding of the text. This tendency to compare different translations to arrive at an understanding has likewise been reported earlier. For example, Tinsley et al (2012)

and Gao et al (2015) discuss how users of unedited MT compare different MT outputs, while Pym & Matsushita (2018) report on users comparing human translations. In the patent case, some tool manufacturers facilitate access to alternative MT engines by embedding links to them directly in the tools. The most common alternative tools that informants reported using were the MT engines embedded in the patent databases of individual country patent offices, such as J-PlatPat in Japan and the Patent Search and Analysis tool in China, but Google Translate was also used occasionally. [3]

5.4. *Larger network of stakeholders*

Beyond the dyad with technical experts, the network that patent professionals rely on includes many other human stakeholders, all of whom have some level of interest in the IPR case in which a machine-translated patent document is being used. In the patenting process, one or more IPR service providers may be employed during the initial drafting of the application, and at least one governmental patent office will examine and prosecute the patent application. Furthermore, the IPR world is highly international and it is rare that a patent would be applied for in only one country. For this reason, further IPR service providers in other countries are often employed to interact with further governmental patent offices, frequently with one of the service providers coordinating the activities. The result is a large and complex network of people and long chains of communication. However, it also means that patent professionals have people they can rely on for help in their task of understanding an L3 patent document. If a patent search results in a document with only one detail that is not understood, patent professionals often tap into this network of people for help instead of ordering a human translation for that part of the document. One of the service providers from the same country the relevant document originated from can be asked to explain the part that was not understood.

Due to the nature of IPR work, even competitors can assume a role in helping to achieve an understanding of a patent document. The patenting process includes a nine-month phase for third-party observation and opposition. According to the Finnish Patent and Registration Office (PRH 2018, 19), the amount of material to review in a patent examination is so large that it would be impossible to cover it all. The purpose of this nine-month phase is therefore to allow third parties to identify relevant cases and information that might have been overlooked during the examination. During this phase, competitors can also challenge a patent and its claims. This often leads to discussions about differing interpretations of claims by different parties. If the case involves relevant patent documents that were machine-translated, those are discussed along with all other relevant documents.

Challenges to the interpretation of MT output might also arise, such as in a case brought up by one of the study informants in which they challenged the MT output used by a national patent office with a translation they took from the European Patent Office's MT engine, which was commonly considered to be a better one. Another informant mentioned a situation in which they challenged the scope of a patent document by offering more information than was contained in the (required) English abstract with a machine translation of a more specific part of the full document. In both cases, the discussions led to a better understanding of the raw MT output, so that they are both good examples of how meaning-making in this MT use case often comes through negotiation.

5.5. *Meaning-making through negotiation on a higher level*

An important point about meaning-making through distributed cognition described in this section is that it is embedded in an environment where meaning-making through negotiation is a common practice. Even when all information is in English, IPR texts tend to be dynamic. Applications change during the long patenting process and, more importantly, the meaning of the texts is interpreted, debated, and challenged by different parties. This offers two important supports for the use of raw MT. First, the space allocated in IPR processes for discussions on

[3] J-PlatPat https://www.j-platpat.inpit.go.jp/\ Patent Search and Analysis http://pss-system .cnipa.gov.cn/sipopublicsearch/inportal/i18n.shtml?params=902F004CA61084A284089435EA AEA94F59CC921A916ADCEB

interpretation and meaning also gives space for machine-translated texts to be examined and errors to be spotted. Second, the people working in this environment can be assumed to be accustomed to the idea of meaning coming through negotiation. The assurance that MT output can be challenged and corrected, and the familiarity with the idea of meaning-making through negotiation, may both contribute to making it easier for patent professionals to accept the practice of relying on raw MT in important processes.

6. Conclusions and future research

The first goal of this case study on the use of raw MT by patent professionals was to analyze whether distributed cognition is a useful theoretical lens through which to examine patent professionals' use of raw MT. The second goal was to contribute to our understanding of the nature of MT gisting through an examination of the environment surrounding a functioning use case.

The study revealed that patent professionals perform their work as part of a large network of artifacts and people, and that the task of understanding machine-translated texts is often achieved through distributed cognition. At the core, the patent professional and MT engine work together as a *coupled system,* where each component in the system plays an active role in cognition. This is augmented through artifacts such as the original source document and alternative MTs, and through interaction with people, including the technical experts, a large network of other stakeholders, and even competitors. The concept of distributed cognition proved to be a useful framework to analyze and explain the understanding of raw MT in this case, and it might be applicable to other cases of MT gisting as well.

The article fulfilled the second goal by providing a nuanced account of the use of raw MT in a real-life scenario. Analyzing patent professionals' MT gisting as a case of distributed cognition expanded our overall understanding of raw MT use and contributed to our knowledge of the conditions that are present in the environment of a functioning case of MT gisting.

There were, however, certain limitations to this case study. First, the study used only one data gathering method, interviewing. This was caused by feedback from informants and potential informants that they could not afford to participate in more time-consuming processes. As discussed by Risku (2010) and others, studies in this area can include a variety of approaches and methods. A more comprehensive picture of distributed cognition could be gained through further studies which employ different methods, for example, surveys would be a way to gather information from more participants and it would not require much time from them.

A second issue is with interviewing itself, which results in a description of informants' ideas of what they are doing, but is not an actual account of their true actions. However, participants' viewpoints on what they are doing is a good starting point, and further studies can build on and test the ideas that come from such a study. A final limitation was that, although one of the main motivations for the study was to examine the role of environment, it was not planned specifically to be an application of a distributed cognition perspective. Perhaps if it had been initially designed as such, it might have revealed even further aspects and nuances regarding the distributed nature of patent MT cognition.

Further study of this use case and this professional group is essential. Their use of raw MT is a widespread and long-term practice, and there is much we can learn about the ways people have found to use raw MT, their strategies to ensure understanding, and their attitudes towards the practice. The knowledge gained from the use case may help us understand current and future uses of MT and other forms of artificial intelligence. Future studies could include a larger number as well as a wider diversity of patent professional users. It would also be beneficial to conduct studies employing varying methods so as to develop a more comprehensive understanding of the phenomenon.

Also, this article's focus on distributed cognition was not an exhaustive account of the different aspects of cognition in relation to this use case; future studies could explore aspects such as embodied or embedded cognition, which would contribute new viewpoints to the phenomenon. More studies would be welcome that focus on other professional environments where raw MT is being used, for example, in academia or customer support scenarios. That might provide new insights into the idea of raw MT as distributed cognition while also building

on our knowledge of the environments and conditions in which MT gisting occurs and the factors that contribute to successful use of raw MT. Most of all, it is evident that raw MT use is a largely under-researched area and is in need of additional studies of all kinds.

References

Alberts, Doreen, Cynthia Barcelon Yang, Ken Koubek, Suzanne Robins, Matthew Rodgers, Edlyn Simmons, and Dominic DeMarco. 2017. "Introduction to Patent Searching." In *Current Challenges in Patent Information Retrieval*. 2nd ed. Edited by M. Lupu, K. Mayer, N. Kando and A. J. Trippe, 3–45. Berlin: Springer.

Battistelli, Benoît. "Patent Translate Service Goes Live!" https://blog.epo.org/the-epo/patent-translate-service-goes-live/ Accessed 5 July 2019.

Bowker, Lynne, and Jairo Buitrago Ciro. 2019. *Machine Translation and Global Research: Towards Improved Machine Translation Literacy in the Scholarly Community*. Bingley, UK: Emerald Publishing.

Bowker, Lynne, and Melissa Ehgoetz. 2007. "Exploring User Acceptance of Machine Translation Output: A recipient evaluation." In *Across Boundaries: International perspectives on Translation Studies*. Edited by D. Kenny and K. Ryou, 209–224. Newcastle, UK: Cambridge Scholars.

Braun, Virginia, and Victoria Clarke. 2006. "Using Thematic Analysis in Psychology." *Qualitative Research in Psychology* 3 (2): 77–101.

Braun, Virginia, and Victoria Clarke. 2013. *Successful Qualitative Research: A practical guide for beginners*. Los Angeles: SAGE.

Burgett, Will. 2015. "Unmoderated Remote Usability Testing of Machine Translation Content." *TAUS Review of Language Business and Technology* IV: 30–37.

Castilho, Sheila, Fabio Alves, Sharon O'Brien, and Morgan O'Brien. 2014. "Does Post-Editing Increase Usability? A study with Brazilian Portuguese as target language." In *Proceedings of the 17th Annual Conference of the European Association for Machine Translation*, 183–190. http://www.mt-archive.info/10/EAMT-2014-Castilho.pdf

Castilho, Sheila, and Sharon O'Brien. 2017. "Acceptability of Machine-Translated Content: A multi-language evaluation by translators and end-users." *Linguistica Antverpiensia* 16: 120–136.

Cavalier, Tim. 2001. "Perspectives on Machine Translation of Patent Information." *World Patent Information* 23 (4).

Clark, Andy, and David Chalmers. 1998. "The Extended Mind." *Analysis* 4 (1): 7–19.

Doherty, Stephen, and Sharon O'Brien. 2009. "Can MT Output be Evaluated through Eye Tracking?" In *Proceedings of the 12th Machine Translation Summit*, 1–8. http://www.mt-archive.info/MTS-2009-Doherty.pdf

Doherty, Stephen, and Sharon O'Brien. 2014. "Assessing the Usability of Raw Machine Translated Output: A user-centered study using eye tracking." *International Journal of Human-Computer Interaction* 30: 40–51.

Ehrensberger-Dow, Maureen. 2019. "Ergonomics and the Translation Process." *Slovo.Ru: Baltijskij Accent* 10 (1): 37–51. https://doi.10.5922/2225-5346-2019-1-3.

EPO: European Patent Office. 2018. *How to Get a European Patent: Guide for applicants*. Vienna: European Patent Office Vienna.

Fuji, M., N. Hatanaka, E. Ito, S. Kamei, H. Kumai, T. Sukehiro, T. Yoshimi, and H. Isahara. 2001. "Evaluation Method for Determining Groups of Users Who Find MT "Useful". In *Proceedings of the 8th Machine Translation Summit*, 11–21. http://www.eamt.org/events/summitVIII/papers.html

Gao, Ge, Hao-Chuan Wang, Dan Cosley, and Susan R. Fussell. 2013. "Same Translation but Different Experience: The effects of highlighting on machine-translated conversations." 27 April–2 May 2013.

Gao, Ge, Bin Xu, David Hau, Zheng Yao, Dan Cosley, and Susan R. Fussell. 2015. "Two is Better than One: Improving multilingual collaboration by giving two machine translation outputs."14–18 March 2015.

Gaspari, Federico. 2004. "Online MT Services and Real Users' Needs: An empirical usability evaluation." In *Machine Translation: From Real Users to Research 6th Conference of the Association for Machine Translation in the Americas, AMTA 2014*, 74–85. https://link.springer.com/book/10.1007/b100780

Gaspari, Federico. 2006. "The Added Value of Free Online MT Services: Confidence boosters for linguistically-challenged Internet users, a case study for the language pair Italian-English." In *Proceedings of the 7th Conference of the Association for Machine Translation of the Americas*, 46–55. http://www.mt-archive.info/05/AMTA-2006-TOC.htm

Gaspari, Federico. 2007. "The Role of Online MT in Webpage Translation." PhD dissertation, University of Manchester.

Henisz-Dostert, Bozena. 1979. “Users' Evaluation of Machine Translation.” In *Machine Translation*. Edited by W. Winter, 149–244. The Hague: Mouton.

Ishida, Toru. 2016. “Intercultural Collaboration and Support Systems: A brief history.” In *PRIMA 2016: Principles and Practice of Multi-Agent Systems, 19th Annual Conference,* 3–19. https://link.springer.com/book/10.1007/978-3-319-44832-9

Joho, Hideo, Leif Azzopardi, and Wim Vandervauwheded. 2010. “A Survey of Patent Users: An Analysis of Tasks, Behavior, Search Functionality and System Requirements.” In *Proceedings of the Third Symposium on Information Interaction in Context: IIiX '10*, 13–24. https://dl.acm.org/citation.cfm?id=1840784.1840789

Jolley, Jason R., and Luciane Maimone. 2015. “Free Online Machine Translation: Use and perceptions by Spanish students and instructors.” In *Learn Languages, Explore Cultures, Transform Lives*. Edited by A. J. Moeller, 181–200. Minneapolis, MN: Central States Conference on the Teaching of Foreign Languages.

List, Jane. 2012. “Review of Machine Translation in Patents - Implications for search.” *World Patent Information* 34 (3): 193–195.

Liu, Nancy Xiuzhi, and Matthew Watts. 2019. “Mobile Translation Experience: Current state and future directions.” In *Impacts of Mobile use and Experience on Contemporary Society*. Edited by X. Xu, 193–212. Hershey, PA: IGI Global.

Marzouk, Shaimaa, and Silvia Hansen-Schirra. 2019. “Evaluation of the Impact of Controlled Language on Neural Machine Translation Compared to Other MT Architectures.” *Machine Translation* 33 (1-2): 179–203. https://doi.org/10.1007/s10590-019-09233-w

Moorkens, Joss. 2017. “Under Pressure: Translation in times of austerity.” *Perspectives: Studies in Translation Theory and Practice* 25 (3): 464–477.

Morland, D. Verne. 2002. *Nutzlos, Bien Pratique, Or Muy Util? Business users speak out on the value of pure machine translation*. Dayton, OH: NCR Corporation.

Muñoz Martin, Ricardo. 2010. “Leave no Stone Unturned: On the development of Cognitive Translatology.” *Translation and Interpreting Studies* 5 (2): 145–162.

Muñoz Martin, Ricardo. 2017. “Looking Toward the Future of Cognitive Translation Studies.” In *The Handbook of Translation and Cognition*. Edited by J. W. Schwieter and A. Ferreira, 555–572. London: Wiley-Blackwell.

Nitzke, Jean, Silvia Hansen–Schirra, and Carmen Canfora. 2019. “Risk Management and Post-Editing Competence.” *The Journal of Specialised Translation* (31): 239–259.

Nurminen, Mary. 2016. “Machine Translation-Mediated Interviewing in Qualitative Research: A Pilot Project.” *New Horizons in Translation Research and Education* 4: 66–84.

Nurminen, Mary, and Niko Papula. 2018. “Gist MT Users: A Snapshot of the Use and Users of One Online MT Tool.” *Proceedings of the 21st Annual Conference of the European Association for Machine Translation,* 199–208. http://rua.ua.es/dspace/handle/10045/76049

Nurminen, Mary. 2019. “Decision-making, Risk, and Gist Machine Translation in the Work of Patent Professionals.” *Proceedings of the 8th Workshop on Patent and Scientific Literature Translation*, 32–42. https://www.aclweb.org/anthology/W19-7204/

Nuutila, Pertti. 2005. “Rough Machine Translation in the Communication Process.” Licentiate thesis, University of Tampere.

O'Brien, Sharon. 2010. “Controlled Language and Readability.” In *Translation and Cognition*. Edited by G. M. Shreve and E. Angelone, 143–165. Amsterdam: John Benjamins.

O'Brien, Sharon. 2017. “Machine Translation and Cognition.” In *The Handbook of Translation and Cognition*. Edited by J. W. Schwieter and A. Ferreira, 313–331. London: Wiley-Blackwell.

Oesch, Rainer, Heli Pihlajamaa, and Sami Sunila. 2014. *Patenttioikeus* [Patent Law]. 3rd edition. Helsinki: Talentum.

Ogura, Kentaro, Yoshihiko Hayashi, Saeko Nomura, and Toru Ishida. 2004. “User Adaptation in MT-Mediated Communication.” *The First International Joint Conference on Natural Language Processing (IJCNLP-04)*, 596–601.

Pituxcoosuuvarn, Mondheera, Toru Ishida, Naomi Yamashita, Toshiyuki Takasaki, and Yumiko Mori. 2018. “Machine Translation Usage in a Children's Workshop.” *CollabTech 2018: Collaboration Technologies and Social Computing: 10th International Conference, Proceedings*, 59–73. Edited by Takashi Yoshino et al, 59–73. Saskatoon, SK, Canada: Springer.

Pouliquen, Bruno. 2016. *Practical use of Machine Translation in International Organizations*. Varanasi, India: World Intellectual Property Organization (WIPO).

PRH: Patentti- ja rekisterihallitus [Finnish Patent and Registration Office]. 2018. Patenttiopas [Patent Guide]. Helsinki: Patentti- ja rekisterihallitus (PRH).

Pym, Anthony, and Kayo Matsushita. 2018. “Risk Mitigation in Translator Decisions.” *Across Languages and Cultures* 19 (1): 1–18.
Risku, Hanna. 2010. “A Cognitive Scientific View on Technical Communication and Translation: Do embodiment and situatedness really make a difference?” *Target* 22 (1): 94–111.
Risku, Hanna. 2014. “Translation Process Research as Interaction Research: From mental to socio-cognitive processes.” In *Minding Translation*. Edited by R. Muñoz, 331–353. *MonTI*, special issue 1. Alicante: Universidad de Alicante.
Rossi, Laura, and Dion Wiggins. 2013. “Applicability and Application of Machine Translation Quality Metrics in the Patent Field.” *World Patent Information* 35 (2): 115–125. http://doi.org/10.1016/j.wpi.2012.12.001.
Schwieter, John W., and Aline Ferreira, eds. 2017. *The Handbook of Translation and Cognition*. Hoboken, NJ: John Wiley & Sons.
Shreve, Gregory M., and Erik Angelone, eds. 2010. *Translation and Cognition*. Amsterdam: John Benjamins.
Smith, Ross. 2003. “Overview of PwC/Sytranet On-Line MT Facility.” In *Proceedings of the Twenty-fifth International Conference on Translating and the Computer*. http://mt-archive.info/Aslib-2003-TOC.htm
Stewart, Osamuyimen, David Lubensky, Scott Macdonald, and Julie Marcotte. 2010. “Using Machine Translation for the Localization of Electronic Support Content: Evaluating end-user satisfaction.” In *Proceedings of the 9th Conference of the Association for Machine Translation in the Americas*. http://www.mt-archive.info/10/AMTA-2010-TOC.htm
Tinsley, John, Alexandru Ceausu, Jian Zhang, Heidi Depraetere, and Joeri Van de Walle. 2012. “IPTranslator: Facilitating patent search with machine translation.” 28 October–1 November, 2012.
Turovsky, Barak. 2016. *Google Translate Blog: Ten years of Google Translate*. https://Blog.Google/Products/Translate/Ten-Years-of-Google-Translate/. Accessed 30 June, 2016.
Wang, Dan. 2009. “Chinese to English Automatic Patent Machine Translation at SIPO.” *World Patent Information* 31 (2): 137–139.
Way, Andy. 2013. “Traditional and Emerging Use-Cases for Machine Translation.” Paper presented at Translating and the Computer 34. https://www.computing.dcu.ie/~away/PUBS/2013/Way_ASLIB_2013.pdf
Way, Andy. 2018. “Quality Expectations of Machine Translation.” In *Translation Quality Assessment: From principles to practice*. Edited by J.Moorkens, S. Castilho, F. Gaspari and S. Doherty, 159–178. Cham, Switzerland: Springer.
WIPO: World Intellectual Property Organization. 2018. *World Intellectual Property Indicators 2018*. Geneva: World Intellectual Property Organization.
Xu, Bin, Ge Gao, Susan R. Fussell, and Dan Cosley. 2014. “Improving Machine Translation by Showing Two Outputs.” In *Chi '14: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 3743–3746. https://dl.acm.org/citation.cfm?id=2556288
Yamashita, Naomi, Rieko Inaba, Hideaki Kuzuoka, and Toru Ishida. 2009. “Difficulties in Establishing Common Ground in Multiparty Groups using Machine Translation.” In *CHI '09: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 679–688. https://dl.acm.org/citation.cfm?id=1518701
Yang, Jin, and Elke Lange. 2003. “Going Live on the Internet.” In *Computers and Translation: A Translator's Guide*. Edited by H. Somers, 191–210. Amsterdam: John Benjamins.
Yasouka, Mika, and Pernille Bjorn. 2011. “Machine Translation Effects on Communication: What makes it difficult to communicate through machine translation?” In *Proceedings of the Second International Conference on Culture and* Computing, 110–115.

Author's address

Mary Nurminen
Faculty of Information Technology and Communication Sciences
Tampere University
Kalevantie 4
FI-33100 Tampere
Finland

mary.nurminen@tuni.fi

Biographical note

Mary Nurminen is currently completeing her PhD on non-translator users of machine translation. She is a teacher of Translation Studies at Tampere University. Before joining academia in 2013, Mary spent 13 years at Nokia and 5 at Lionbridge, working in the areas of technical communication and localization as (in this order) practitioner, manager, and solution architect.

PUBLICATION IV

Machine Translation-Mediated Interviewing as a Method for Gathering Data in Qualitative Research: A Pilot Project

Mary Nurminen

New Horizons in Translation Research and Education 4(16): 66–84

TURO RAUTAOJA, TAMARA MIKOLIČ JUŽNIČ AND KAISA KOSKINEN (EDS)

New Horizons in Translation Research and Education 4

Publications of the University of Eastern Finland
Reports and Studies in Education, Humanities and Theology
No 16

University of Eastern Finland
Joensuu
2016

Grano Oy
Joensuu, 2016
Editor-in-chief: Maija Könönen
Sales: Itä-Suomen yliopiston kirjasto

ISBN: 978-952-61-2356-1 (PDF)
ISSN: 1798-565X

Machine Translation-Mediated Interviewing as a Method for Gathering Data in Qualitative Research: a Pilot Project

Mary Nurminen, University of Tampere, Finland

ABSTRACT

This article describes a project in which machine translation (MT)-mediated interviewing was used to gather data on the end users of an online application for machine translating PDFs. Four interviews with Spanish speakers were implemented using Skype Translator's instant messaging (IM) function as a medium for communication. Seven considerations on the method that arose in the project are discussed. Two of these concerned the use of IM as a medium for interviewing, namely, considerations of time zones and multitasking on the part of the interviewees. Five considerations arose that were centered specifically on MT-mediated interviewing: technology, time requirements, understanding and negotiation for meaning, participants' target language knowledge and adaptation, and user experience. These considerations can be seen as the beginning of a definition of best practices for MT-mediated interviewing.

KEY WORDS: machine translation, machine translation for communication, MT-mediated communication, MT-mediated interviewing, instant messaging

1 INTRODUCTION

I had already been considering the idea of using machine translation (MT)-mediated communication as a method for gathering research data when I started discussions with the Finnish company Multilizer in the summer of 2015. They explained that were interested

in learning more about the users of an internet-based tool they manufacture, PDF Translator.[1] They understood something about how the tool was being used through automatically generated statistics, and they understood something about their users through the web-based questionnaires they held occasionally. However, they were interested in gaining a deeper understanding of their end users and were curious about research methods that could lead them to that understanding. The result of the discussions was the launch of a cooperative project with two goals:

Goal 1: piloting the use of MT-mediated interviewing as a research method
Goal 2: gathering data on the end users of Multilizer's PDF Translator tool

The expected results for goal 1 were that the method would prove to be promising enough to warrant further study and testing, and that some factors would be revealed which can affect the use of the method. The expected results for goal 2 were that the interviews would uncover new information about PDF Translator users.

This paper focuses on the results of goal 1, the piloting of the use of MT-mediated interviewing as a data-gathering tool. The results of the goal 2 were communicated to Multilizer in a final project report in March 2016 and are not in the scope of this study.

1.1 MT-mediated interviewing

Several factors in Multilizer's situation indicated that interviewing would be a good method for gaining the understanding they were looking for. First, the focus of the interview would be an internet-based tool and its usability. As Jakob Nielsen states, "Many aspects of usability can best be studied by simply asking the users." (Nielsen 1993:209) Second, the information Multilizer would receive would be combined with information already gathered through other methods to construct a more holistic picture of users (Hirsjärvi & Hurme 2011).

As explained earlier, I had an interest in using interviews because I wanted to pilot the idea of interviewing over MT as a method for data gathering. I was aware that MT-mediated communication was already in use in various areas of business, for example by customer service representatives to support customers with whom they do not share a language (Burgett et al. 2012) or in online community forums (Burgett et al. 2012; Mitchell & Roturier 2012). I believed it would be worthwhile to try applying the approach in research.

A search of the literature on interviewing in research did not reveal studies employing MT-enabled interviewing as a method, nor did the literature on MT reveal studies in which MT was used in an interviewing context. It seemed that there was a gap in research on this particular context for using MT-mediated communication. However, both interviewing over instant messaging and MT-mediated communication in other contexts have received increasing attention since the early 2000s.

[1] pdf.translator.com

1.2 Interviewing over Instant Messaging (IM)

One of the data-gathering methods that the internet has made possible is interviewing using instant messaging (IM) applications, and the use of this medium has grown, "particularly…if the research explores an Internet-based activity such as e-learning or online community, where the research participants are already comfortable with online interactions" (Kazmer & Xie 2008:257). Mann and Stewart (2002) point out that interviewing in this context may be more natural to some interviewees than face-to-face interviewing would be.

Several interesting studies (Kazmer & Xie 2008; Opdenakker 2006) compare IM interviewing with methods like interviewing face-to-face, by telephone, or by e-mail. Other studies (Markham 2004; Voida et al. 2004) delve into the features, advantages and disadvantages of IM interviewing itself. Among the advantages of IM interviewing outlined in the studies are access to a very wide array of potential participants and a reduced need to travel to conduct interviews, meaning a reduction in project costs. One very clear advantage is that IM applications normally retain interview data in one file, meaning that no transcribing is needed after interviews, although as Opdenakker (2006) points out, this can lead to a reduction in note-taking, which can be detrimental to results.

The challenges of IM interviewing are also well covered. Both Markham (2004) and Voida et al. (2004) discuss the difficulties of learning to suppress their desire to reply overly quickly to interviewees, an act which can interrupt the interviewees' line of thought and comment. Several researchers (Markham 2004; Opdenakker 2006; Voida et al. 2004) cite the lack of the social cues we are used to relying on in face-to-face communication as potentially detrimental. IM chats are also prone to discontinuities and overlapping messaging, which can cause extra work in the analysis phase. It is interesting to ponder whether these are disadvantages to us now, as we learn to use new forms of communication, but will be so natural to future generations that they will no longer see them as disadvantages but as simple features of communication.

1.3 MT-mediated communication

Hutchins (2010) outlines three main types of use for machine translation (MT), which are described in Table 1.

Table 1: Three main types of MT use (Hutchins 2010)

MT use type	Description
MT for dissemination	Information is put through MT and the resulting 'raw' output is edited by humans in a task known as post-editing. The final result is language of publishing-level quality. The information is then disseminated to readers.
MT for assimilation	Information is put through MT and the resulting 'raw' output is consumed directly by a reader who needs a general understanding of the information, but does not need the information to be grammatically or stylistically of publishing-level quality.
MT for communication (MT-mediated communication)	MT is used in social interchange such as e-mail or instant messaging, allowing people to communicate across language barriers. Again, publishing-level quality is not a requirement for the information.

Of these three, MT for dissemination has the largest amount of research devoted to it, with significant contributions from the field of Translation Studies. In this context, MT is seen as one of the aids available to translators to use in their work, and research has addressed topics such as evaluation of MT quality, translators' roles, and processes. The task of post-editing of MT output is the focus of a number of studies; for a good overview of the research, see Koponen (2016). The use and use cases for MT for assimilation and MT for communication have slowly gained momentum over the past 20 years, and the past 5 years have seen very rapid growth. However, this rapid growth in use has not resulted in a similar rapid growth in research, and the amount of research on those phenomena remains limited.

Although the amount of research remains small, MT-mediated communication has been studied since at least 2002, when the Intercultural Collaboration Experiment (ICE) was established between several Asian universities to provide communication tools for multilingual online meetings and collaboration (Nomura et al. 2003). In conjunction with ICE, various aspects of MT-mediated communication were studied and reported on (Nomura et al. 2003; Ogura et al. 2004). Since then, similar studies have been done involving other environments where multilingual communication took place via MT (Yamashita & Ishinda 2006; Yasouka & Björn 2011; Calefato et al. 2012; Gao et al. 2013). Most often these involve experiments with university students as participants. They are placed in geographically distributed, multilingual work teams and given a specific task to complete, with communication related to the task taking place in an online communication tool with embedded MT. Then various aspects of the communication are analyzed.

Calefato et al. (2012) examined how the activeness of participation in discussions was affected when people use their native language over MT instead of English. They found that discussions were more balanced when MT allowed people to use their own languages. In the experiment covered by Ogura et al. (2004), participants wrote messages in their own language, reviewed the MT output in English, and then had a chance to make changes before that output was machine translated further into the languages of their other team members. The study analyzed the types of adaptations they made in their source text messages to produce better MT output in English.

Yamashita and Ishida (2006) looked at how communicators used referring expressions when discussing their tasks, and how machine translation affected the use and success of reference communication. Yasouka and Björn (2011) studied the importance of establishing and maintaining common ground, specifically through techniques such as using project-specific jargon, to the communication process. Their findings indicate that this establishment of common ground plays as important a role as the linguistic quality of the MT in successful MT-mediated communications.

An interesting study by Gao et al. (2013) analyzed how participants' belief in whether MT was in use or not affected their view of the communication experience. Participants were paired up and given a task that they discussed in an online chat. The discussions were in English, although the English-speaking participants did not know whether the messages they received were typed by their Chinese-speaking partners or put through MT. The results showed that the belief of MT being present had a positive effect on the participants' view of the communication experience, perhaps because they could attribute mistakes or ungrammatical language to the machine.

2 THE PROJECT

The project was conducted in July and August of 2015 and comprised interviews with four users of PDF Translator. PDF Translator takes a PDF file, extracts the text, puts the text through machine translation to translate it, re-assembles the file to match the original PDF, and creates a new PDF in the machine-translated language. It is used by people who have a document they want or need to understand, but they do not know the language it is written in. It is therefore a tool enabling MT for assimilation. PDF Translator is available by download in the internet and has a free version that can automatically translate a limited number of pages of text. The paid versions of the tool involve purchasing a 'quota', which is a pre-defined number of pages that users can translate with the tool. The user base of PDF Translator is large - a significant number of new downloads of the free version are completed every day - and diverse, with users across the globe who access any of the 27 languages available.

2.1 The technology

It was assumed that the target audience of the study, users of the MT tool PDF Translator, might be open to participating in an innovative interviewing method that also relies on MT. However, since PDF Translator is an MT tool for assimilation, not communication, a different MT tool would be used for interviewing.

Skype Translator preview was selected as the interviewing tool for several reasons. First and foremost, Skype is widely available and included in many software packages, meaning that it would be easier to recruit participants who already had the technology available. Also, because Skype uses Microsoft's Bing Translator, the quality of the MT for the language pair to be used (English-Spanish) could be assumed to be of good enough quality to support this type of pilot project.

Another decision was to conduct the interviews using the instant messaging function of Skype Translator instead of the video and voice function. Due to Skype's background as

a video and voice tool, as well as the recent press on Skype Translator, which features video and voice, Skype Translator is mostly seen as a tool for spoken language. However, it is also equipped with a text-based IM chat that uses the same MT backbone (Bing Translator) as the video and voice function.

The IM function was chosen for four reasons. The first is that it poses fewer technology requirements for both the interviewer and interviewees. It was assumed that most potential participants already had the technology needed for IM conversations, whereas Skype video and voice calls require not only a computer and very solid internet connection, but also a camera and voice equipment. The second reason was that the IM involves a simpler technology with fewer components that need to communicate with each other to produce good results, meaning that it would be less likely to have problems. A third reason focused on the participants: people who are not familiar with video calling may feel uncomfortable using it in an interview situation. The final reason for selecting IM was that no transcription of the interviews would be needed. Once the interviews were over, the transcription of the conversations would be ready. As mentioned in the literature on IM interviewing (e.g. Opdenakker 2006), this has been cited as a considerable advantage.

At the time of the interviews, Skype Translator was available in a preview version and was separate from the traditional Skype application. The former had to be downloaded separately and had more strict technical requirements than Skype. However, for bilingual conversations, it was sufficient if one of the participants had the Skype Translator application. The second participant could be working on a regular Skype application, but had the same MT benefits as the Skype Translator participant.

During Skype Translator chatting, each participant enters their text in their own language. The application translates that text and can be configured to show both the original and the machine translated text to each participant, with their own language always shown at the top. The following example shows an excerpt from an anonymized interview. This excerpt was taken directly from Skype Translator to highlight the view the user has while working.

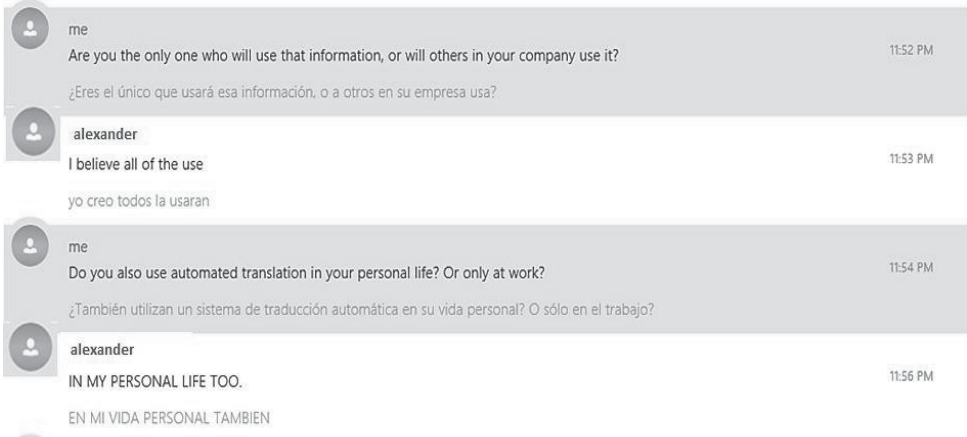


Figure 1: Skype Translator chat, view the user has while working

2.2 Recruiting interviewees

As I was inexperienced in using MT-mediated communication with speakers of languages other than English, I decided to limit the recruitment of interviewees to only those users who downloaded the Spanish-language version of PDF Translator. I have a basic understanding of Spanish and I anticipated that it might be helpful to be able to read the source texts. I took into consideration that partial knowledge of the language could affect the outcome and would make it inherently different from MT-mediated communication that involves participants who have no knowledge of each other's languages, but decided that this would be the best approach to ensure the success of both goals of the project. I return to this issue in the Discussion section.

Another reason for selecting Spanish speakers was that the Spanish-English language pair is often recognized as one of the most favorable ones for MT. An example of this is the maturity check conducted by the European Commission in 2011. This evaluated the MT results for 21 languages paired with English, and Spanish was found to be the one that produced the best results (Reiman 2014).

Interviewees for the study were recruited through a short questionnaire that was displayed to all PDF Translator users who downloaded the Spanish language version of the tool. This questionnaire requested information on e-mail addresses, Skype names, willingness to be interviewed, and a question regarding the type of information that they used PDF Translator to translate. A reward of 100 pages of free translation quota was offered to all who participated in the interviews. At the top of the recruiting questionnaire was a statement that the information collected was for a research project and would not be used for any other purpose than this specific project. Later in the interviews, it was again explained that the information would be used for research purposes only and that all participants would be anonymous.

Initially I used e-mail to contact people for scheduling Skype IM interviews. I soon noticed that the response rate for this was very low: out of 15 invitations sent, I received only 1 response. Over the course of the ensuing e-mail conversation to schedule that one interview, the person quit responding. I decided to change tactics and I began to send invitations to users directly in Skype. This proved to be a more effective solution. I eventually recruited and conducted full interviews with four users. In all four cases, a key factor in successfully recruiting interviewees was catching the person online in Skype in real time. Once synchronous communication was established, all four were able to begin the interview immediately or within 30 minutes.

The interviewees were all male, between the ages of 38 and 52, and all had either a technician or university-level degree. Two had an educational background in computer science or information technologies, a third reported his proficiency with computers to be "100%" and the fourth reported average computer skills. None had broad competence in any language other than their native Spanish. Three reported having some knowledge of English, which they described as "a little", "very little", and "low". One reported having no knowledge of English. All were located in Central and South America.

2.3 The interviews

Semi-structured interviewing was chosen because we wanted to get comparable information on certain themes from the four interviewees but at the same time leave flexibility to ask follow-up questions or move to topics brought up by the interviewees (Hirsjärvi & Hurme 2011). Flexibility might also be needed because of the nature of MT-mediated communication, which might require additional questions. The focus of the interviews was the interviewees' use of PDF Translator, and the majority of the questions centered on themes around that, with the aims of both gathering information for goal 2 of the project and act as the pilot for goal 1. At the end of the interview, one question was asked which focused specifically on goal 1: what was the experience of being interviewed via MT like for the interviewees. The themes covered are shown in Table 2 below.

Table 2: Themes covered in the interviews

Theme	Description
Initial data gathering	• General questions about the interviewee: age, level of education, subject of degree, current profession, level of proficiency in languages and use of computers.
About the translation they got from PDF Translator	• Material they translated: genre (type), where it was obtained, what would be done with the information, how many people would use the information • Perceived quality: what was the user's overall impression of the translation quality, what expectations did they have for the translation, and how well those expectations were met
About the use of PDF Translator	• Perceived ease of use of the tool: how long it took them to install it and get their translation • Other needs and tools used: how often they have the need to translate documents, what other tools they use for that, what languages were involved • Ideas for the tool: other things they hoped the tool would be able to do
Wrap-up	• Any further information they wanted like to give about PDF Translator. • Questions about the interview experience: how well they think MT worked, did they feel they were understood, and would they recommend this method of communication to their friends • Reminder that the information gathered was for research purposes only and that they would remain anonymous (either at beginning or end of interview)

The interviews were scheduled to be 30 minutes but lasted longer. The shortest was 42 minutes and the longest was 73 minutes. This was necessary to cover all of the questions I intended to ask, but also for the extra clarification requests and negotiation of meaning that is needed in MT communication. The timing did not afford much opportunity for establishing rapport or branching off into other areas that arose in our conversations.

2.4 Data compilation

As discussed in the introduction, one benefit of interviewing over IM is that the researcher does not need to transcribe audio files prior to starting their analysis. Skype keeps all interactions between two IM participants in one file, which is easy to download or copy/paste into another format for further processing. Even when the communication

includes asynchronous messages spread out over several days, such as during the interview-scheduling phase, the messages are saved in one file. In this project I transferred the data into Word, anonymized it by replacing interviewee names with pseudonyms, then formatted it to facilitate analysis. This was a simple operation and a time saver for me. Due to the study restriction in the number of interviews, the data was not transferred to a qualitative data analysis tool, but the transfer would likely have been a simple operation. An overview of the data gathered from the interviews is given in Table 3 below.

Table 3: Overview of interviews

Interview	Time (minutes)	Total word count	Number of turns	Number of unique questions asked
1	69	1529	84	31
2	50	2201	77	32
3	42	1611	70	30
4	73	1247	55	27

As with other interview types, some interviews involved more "talk" than others, although a somewhat comparable number of unique questions was covered. There was some variance in the length of the interviews, which seems to have no correlation with the number of speaking turns taken or the number of unique questions covered. This reflects the variance in how focused participants were on the interview: while some appeared to be concentrating exclusively on the interview, others seemed to be multitasking. I return to this in the Discussion.

3 DISCUSSION

My conclusion from this small pilot project was that MT-mediated interviewing is a data-gathering method worth further exploration. The pilot revealed some important considerations for using MT-mediated interviewing which could be helpful to other researchers who consider using the method. They could also be the start of an eventual understanding on best practices for using the method.

Seven considerations arose from the pilot project. Two of these, considerations of time zones and multitasking, are aspects that apply specifically to interviewing over IM, and they would be the same whether those interviews had been conducted between speakers of the same language or between speakers who were communicating through MT. In fact, my findings on time zones and multitasking reflect the results in studies on unilingual IM interviewing (Kazmer & Xie 2008; Voida et al. 2004). The other five considerations apply specifically to MT-mediated interviewing and include considerations of technology, time requirements, understanding and negotiation for meaning, participants' target language knowledge and adaptation, and user experience.

3.1 Considerations concerning IM interviewing

3.1.1 Time zones

My experience in this project mirrored that of Kazmer and Xie, who reported that "scheduling can be quite complicated especially when the medium is synchronous, in which case two primary factors come into play: time zones and local scheduling conflicts." (Kazmer & Xie 2008:262). I recruited participants from the group of all people who downloaded the Spanish version of PDF Translator during the time frame of the project. This meant that the majority of potential participants were located in North or South America, in time zones eight to nine hours earlier than my own. Although it was not intentional, all four of the eventual recruited participants were from Central and South America. This had two consequences.

The first consequence was that, at the time of the interviews, the interviewees were at work. It is possible that people considered it acceptable to grant an interview at work because in it, they would be discussing a tool that many of them were using at work. Another consideration is that typing in an IM tool does not resemble face-to-face interviewing and this meant that they could participate without being noticed. In fact, one participant remarked that if the interview were to involve video, he would not be able to participate until 8 p.m. because "at work is complicated," whereas if it were an IM interview, he could participate immediately. It is clear that for longer interviews, or interviews covering distinctly personal topics, a more appropriate time for interviewing might be in the evening.

The second factor arising from the time difference was that it was necessary for me as the interviewer to work outside of normal working hours. Through trial and error, I found that it was most effective to establish initial contact, recruit and interview people in the late evening hours of my time zone, requiring that I rearranged my schedule to be available and alert. It was a good reminder that although modern technology can help us overcome many restrictions in research, we still need to plan around certain practical limitations.

3.1.2 Multitasking

In using IM, the interviewer cannot determine whether the interviewee is giving their full attention to the interview, as they would in a traditional face-to-face or telephone interview. They might also be multitasking while also chatting with the interviewer, which would reflect the typical way IM is used. My impression when interviewing was that the interviewees were most likely doing other tasks in addition to chatting with me. However, when reviewing the transcripts, the overwhelming majority of responses came within two minutes of the submission of the previous chat turn. Two of the interviewees exhibited no response lag of greater than two minutes. One interviewee had only one lag of over two minutes. The fourth interview was noticeably different. Although it lasted the longest time, it produced the lowest numbers in total word count, turns, and unique questions asked. It was clear that the interviewee was doing other things while responding. However, that interview was also completed and no significant differences in results were detected. It would seem that, even if participants were performing other tasks in addition to answering

interview questions, those other activities were not so long-lasting or absorbing that they would affect the overall completion of the interviews.

3.2 Considerations Specific to MT-mediated Interviewing

3.2.1 Technology

Many tools are available for MT-mediated text communications. Some are in commercial use in closed environments, such as those used by technical support agents who support customers who speak a different language. Others are freely available on the internet. Aiken et al. (2009) listed eight chat applications integrated with MT (both commercial and free). Such a list changes rapidly and we can assume that there are more applications available today.

Skype Translator's preview version was a suitable platform for this type of interviewing, especially since it only required one of the participants to have the Translator version. As the regular version of Skype was free and readily available globally, it meant there were no overwhelming technical demands for potential interviewees. Furthermore, many people already had Skype installed on their computers, so it required no extra downloading or configuration work on their part. This made the task of recruiting willing participants easier.

Only once during this project did internet connectivity issues interfere in an interview, in the form of a minor and short-lived slowing of the internet. This was detected by both the interviewee and myself, but was brief and was probably caused simply by the wireless infrastructure in my location.

3.2.2 Time requirements

It was clear in the interviews that the 30 minutes I originally allocated for interviews was insufficient. This was a confirmation of Markham's statement on IM interviewing that "Synchronous interviewing online took about twice as long as face to face" (Markham 2004:365). In addition, during MT-mediated interviewing time is also needed to ask for clarification, to adapt texts to produce better translations, and to negotiate meaning. This would indicate that the time required for MT-mediated interviewing is even longer than what Markham suggests. This should be a consideration in planning, and also needs to be communicated to potential participants so that they can suggest an appropriate time for the interview.

This longer time commitment could reduce the number of people willing to be interviewed. The results of this pilot indicate that 45-75 minutes is a time frame people are willing to sacrifice in the middle of their day, at least when there is a small reward offered. However, as the interviews did not continue longer than that, I did not obtain data on the retention rates for longer interviews.

The time commitment required for a longer interview has another negative side in that it makes it more difficult to conduct impromptu interviews. In communities that rely on IM for communication, it is a common practice to "ping" other people, meaning sending them a quick message and seeing if they respond. If they do, an impromptu discussion can ensue. In essence, this is the same tactic I used in recruiting people for this project and it

worked well. As mentioned previously, instead of scheduling an interview for a future time, all participants were willing to start immediately or within half an hour. However, pinging someone to start a discussion and then launching a 90-minute interview might not produce good results. One solution for topics that simply require more time might be recruiting people with the "ping" strategy, then scheduling a short series of 30-minute interviews.

3.2.3 Understanding and negotiation for meaning

When evaluating the possibilities of adopting MT-mediated interviewing for data gathering, one of the main questions concerns whether the communication and understanding in the interview are sufficient to produce reliable data. On the one hand, the idea of gathering data through imperfect communication may seem ill advised. At times during the interviewing, it felt somewhat like working through an interpreter who was somewhat knowledgeable of the terminology of the subject we were discussing, but did not have a good grasp of grammar, and sometimes could not translate a word at all because the speaker did not say it exactly right. The question then arises whether a researcher can claim reliability when there is so much potential for misunderstanding.

On the other hand, interviews inevitably involve factors that potentially hinder understanding. Hirsjärvi and Hurme (2011) discuss the possible effects of participants' different communication styles and levels of linguistic competence – whether those are results of a difference in the participants' social class or simply personal differences. Ruusuvuori and Tiittula (2005) examine interviewing in the light of different situations: when the cultures of the interviewer and interviewee are different, interviewing older people, children, or people with aphasia, and finally, computer-mediated interviewing. Other, smaller factors can affect the interview situation. These include different accents, native and non-native interaction, technical difficulties, even background noises. Even the simple fact of there being two individuals with individual backgrounds, ideas, and understanding of the point of the interview can affect interview outcomes. Yet researchers conduct interviews regardless of all of these factors. To quote a professor of mine when I first asked her about the possibility of using MT-mediated interviewing and the ensuing imperfect language: "Of course we can deal with imperfect language. People do it all the time!"

One available aid we have for increasing and ensuring understanding in spite of imperfect language is simple communication: asking for clarification, repeating, or rephrasing things. My pilot project showed ample evidence of this throughout the interviews, as shown in the excerpts below. Note that the excerpts are taken from my screen and therefore have English on the top and Spanish under it. When I write, the Spanish translation is shown below, whereas when the interviewee, Tomás, writes in Spanish, the translation in English is shown above it. This method allows the reader to follow the conversation easily, focusing mostly on the top text in their own language.

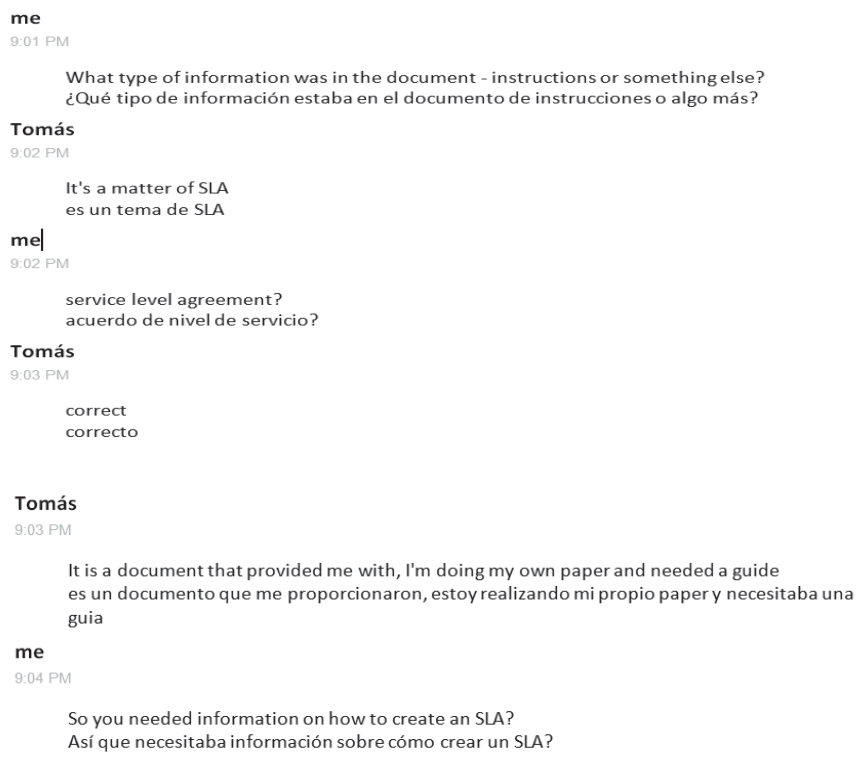


Figure 2: Examples of rephrasing

The first instance of lack of understanding involved an acronym. Fortunately, I happen to know it but I asked for confirmation to make sure. In the second instance, the machine translation was somewhat understandable, but I still needed to make sure I understood so used rephrasing to ask for confirmation.

Many of the gaps in understanding during the interviews were resolved in a similar way. However, not all were clarified and some issues and questions did remain after the interviews had ended. In future studies, it would be advisable to devise methods for overcoming this and ensuring that all necessary information is gathered and understood. One method might be to compile an initial list of questions and have it professionally translated and sent to participants prior to the interview. This would help define the domain and terminology of the conversation. Another idea might be to have a professional translator review the transcripts after the interview, either in full or only for those parts that the interviewer marks for review. This would be more time-efficient and less expensive than employing a translator to conduct or participate in interviews. After the review, interviewees could be contacted for a short follow-up discussion to resolve open issues and questions.

3.2.4 Participants' target language knowledge and adaptation

As mentioned in section 2.2 of this paper, I decided to recruit participants from Spanish-speaking countries because I had a basic understanding of Spanish and thought that that

might be helpful. As it transpired, three of the four interviewees also had some knowledge of English, so in fact we shared the aspect of being able to read the source texts and evaluate their quality to some degree.

The participants' knowledge of the other language surfaced in one very clear way: they tended to adapt source texts to try to produce translations that were more comprehensible to the other party in the interview. This adaptation of the source message reflects the research of Ogura et al. (2004) on the different adaptation strategies used in MT-mediated multilingual conversations. Evidence of this adaptation occurred on the part of both the interviewees and myself. For example, one participant mentioned *pages* several times during the interview. At first he used the Spanish word *hojas,* which was translated into *leaves* in English. In my reply, I used the word *pages,* which was translated as *páginas*.

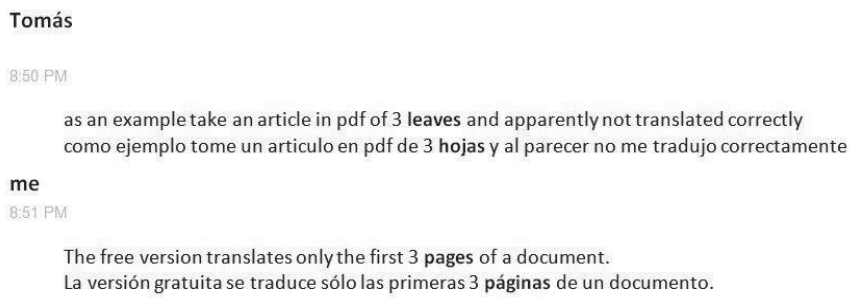


Figure 3. Example of adaptation

Ten minutes later, when we were again discussing pages, the interviewee again used *hojas* but then corrected himself:

Figure 4. Example of adaptation

When the word arose once more 20 minutes later, he again used *página.* He seems to have learned from the MT output that the Spanish word *página* produces a better result in English than the word *hoja* does.

On my own part, I was asked for clarification in interview 3 and successfully changed the verb to produce a better output in Spanish:

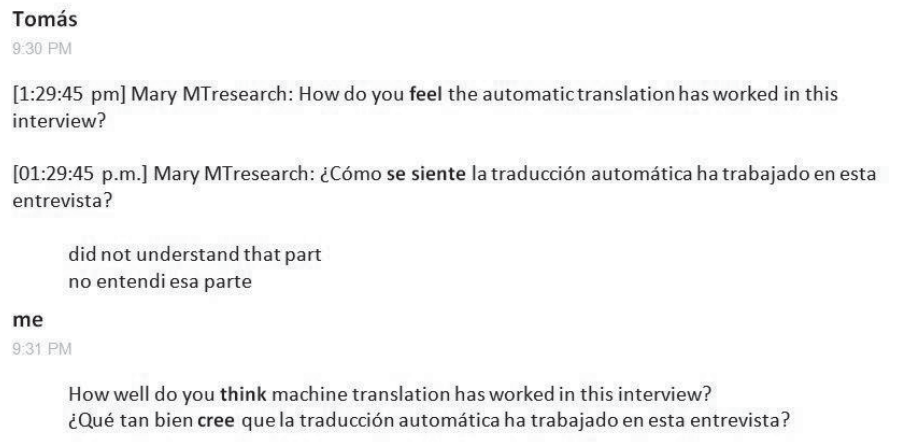


Figure 5. Example of adaptation

These examples indicate that in MT-mediated communication, participants' knowledge of the target language can affect communicative success. Even when participants rate their knowledge of the target language as basic or low, they do seem to be able to use that knowledge to evaluate MT outputs and adapt their messages in hopes of producing better translations.

Another indication that some knowledge of the other language was helpful came in the form of one participant's response to my question, "This interview has been done using automated translation. If your friend asks you about it later, how will you describe your experience?" The participant stated:

> *Very good, because it allows me to review the complete translation in the original language and in my language*
>
> *Muy buena, ya que me permite revisar la traduccion completa en el idioma original y en mi idioma*

Of course, this can only be helpful when participants have access to both source texts and translated texts, not just the translated texts. This is something that the manufacturers of MT and IM applications might want to take into account in their design work.

Another case where user access to both source and translated texts has potential to affect the quality of MT output is when English is used as a pivot language in the MT process. Currently some language pairs are challenging because it is difficult to find enough data to produce good machine translation, so texts might first be translated into English, and the English MT output is then used to translate into the target language. That process is not usually transparent to the end user. They can only guess that that is what is happening based on the MT results they get. However, what if it were made transparent, and MT users were shown the initial translation into English as well as the translation from that into the final target text? In cases where the user knows some English, they would have two texts to rely on for understanding instead of just one. Although this might be

more time-consuming, it could help promote understanding and potentially make the use of MT more successful.

3.2.5 User experience

At the end of each interview, participants were asked about their interview experience and their impressions of MT-mediated communication. All four participants gave a positive or very positive response. I first asked them how they felt the machine translation worked, and responses included expressions such as "well, very good," "understandable," "it all worked," "very good, excellent." When asked if participants felt I had understood everything they had to say or how well they thought we understood each other in the interview, their responses included "yes," "totally," "At 100%, thank you for your attention," and "very clear." I asked two of the interviewees if they would recommend this type of communication to a friend and they responded with "Yes" and "with security" (for sure). As mentioned earlier, the people who volunteered for these interviews represent a portion of the population that is already familiar with digital information and MT, and could be assumed to be more open to working with new technologies. This project shows some indication that, at least with this type of person, the initial experience with MT for communication tends to be positive.

4 CONCLUSIONS

This was a very limited experiment in MT-mediated interviewing. It confirmed some of the observations on interviewing over IM that have been reported in previous studies, and revealed some of the issues to be considered in MT-mediated interviewing.

One of the most interesting findings of the project was the effect of having access to both source and target texts in MT-mediated communication. When participants could see all texts in both languages, even their reportedly low level of knowledge of the target language seemed to be helpful in ensuring successful communications. Another interesting outcome was the participants' positive response to the medium.

The results of the project gave some preliminary indications that MT-mediated interviewing is worth further exploration as a data-gathering method for qualitative research. The most significant benefit of the method is the potential expansion it brings to the size of populations that can be included in research. Studies can be conducted on people who are widely distributed geographically, linguistically and culturally, without an equally large expansion in project resourcing.

The method brings certain challenges with it. Perhaps the largest of these is the potential for misunderstanding, which could lead to questions on reliability and validity. More research on MT-mediated interviewing, and MT-mediated communication in general, could lead to a better understanding of the best practices for using the method. It is hoped that the findings reported on in this article will help to trigger interest in further studies in this area.

Studies comparing this interviewing medium with others, similar to the comparative studies between IM and other types of interviewing by Opdenakker (2006) and Kazmer & Xie (2008), would help to reveal the weaknesses and strengths of the medium, or the

contexts where it is best applied. Another interesting comparison would be between interviews mediated by a human interpreter and those mediated by MT.

In the area of MT-mediated communication, it would be interesting to study the experience of advanced users of the medium. Currently those may be difficult to find, but there is one group that may already qualify: technical support agents in companies that are using MT-mediated communication to offer support in languages their agents do not speak. The experiences of those users could offer valuable input for further research and technology development. In general, it would be good to see more focus on developing methods for evaluating the many issues that can affect the effectiveness of MT-mediated communication.

REFERENCES

Aiken, Milam, Mina Park, Lakisha Simmons, and Tobin Lindblom. 2009. "Automatic Translation in Multilingual Electronic Meetings." *Translation Journal* 13 (3) http://translationjournal.net/journal/49meeting.htm. Visited September 2015.

Burgett, Will, Julie Chang, Ryan Martin, and Yoshi Yamakawa. 2012. *IT@Intel White Paper: Enabling Multilingual Collaboration through Machine Translation*. Intel IT.

Calefato, Fabio, Filippo Lanubile, Tayana Conte, and Rafael Prikladnicki. 2012. "Assessing the Impact of Real-Time Machine Translation on Requirements Meetings: A Replicated Experiment." *Proceedings of the ACM-IEEE international symposium on Empirical software engineering and measurement:* 251-260. Lund, Sweden.

Gao, Ge, Bin Xu, Dan Cosley, and Susan R. Fussell. 2014. "How Beliefs about the Presence of Machine Translation Impact Multilingual Collaborations." *Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing:* 1549-1560. Baltimore, MD, USA.

Hirsjärvi, Sirkka, and Helena Hurme. 2011. *Tutkimus-haastattelu: Teemahaastattelun teoria ja käytäntö* [The Research Interview: Theory and Practice for Focused Interviews]. Helsinki: Gaudeamus Helsinki University Press.

Hutchins, W. John. 2010. "Machine Translation: A Concise History." *Journal of Translation Studies* 13 (1-2): 29-70.

Kazmer, Michelle M., and Bo Xie. 2008. "Qualitative Interviewing in Internet Studies: Playing with the Media, Playing with the Method." *Information, Communication & Society* 11 (2): 257-278.

Koponen, Maarit. 2016. "Is Machine Translation Post-Editing Worth the Effort? A Survey of Research into Post-Editing and Effort." *The Journal of Specialised Translation* 25: 131-148.

Mann, Chris, and Fiona Stewart. 2002. "Internet Interviewing." In *Handbook of Interview Research: Context and Methods,* Jaber F. Gubrium and James A. Holstein (eds), 603-628. Thousand Oaks: Sage.

Markham, Annette M. 2004. "The Internet as Research Context." In *Qualitative Research Practice,* Clive Seale, Giampietro Gobo, Jaber F. Gubrium, and David Silverman (eds), 358-374. London: Sage.

Mitchell, Linda, and Johann Roturier. 2012. "Evaluation of Machine-Translated User Generated Content: A Pilot Study Based on User Ratings." *Proceedings of the 16th Annual Conference for the European Association for Machine Translation*. Trento, Italy.

Nielsen, Jakob. 1993. *Usability Engineering*. London: Academic Press.

Nomura, Saeko, Toru Ishida, and Mika Yasouka. 2003. "Open Source Software Development with Your Mother Language: Intercultural Collaboration Experiment 2002." *10th International Congress on Human-Computer Interaction.* Crete, Greece.

Ogura, Kentaro, Yoshihiko Hayashi, Saeko Nomura, & Toru Ishida. 2004. "User Adaptation in MT-Mediated Communication." *The First International Joint Conference on Natural Language Processing (IJCNLP-04):* 596-601. Hainan Island, China.

Opdenakker, Raymond. 2006. "Advantages and Disadvantages of Four Interview Techniques in Qualitative Research." *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research* 7 (4), Art. 11.

Reiman, Mikael. 2014. "MT@EC: European Commission Machine Translation for Public Administrations in the EU Member States." Paper presented at the annual Kites Symposium, Helsinki, Finland, October 2015.

Ruusuvuori, Johanna, and Liisa Tiittula (eds) 2005. *Haastattelu: tutkimus, tilanteet ja vuorovaikutus* [The Interview: Research, Situations and Communication]. Tampere: Vastapaino.

Voida, Amy, Elizabeth D. Mynatt, Thomas Erickson, and Wendy A. Kellogg. 2004. "Interviewing over instant messaging." *CHI EA '04 CHI '04 Extended Abstracts on Human Factors in Computing Systems*, 1344-1347.

Yamashita, Naomi, and Toru Ishida. 2006. "Effects of Machine Translation on Collaborative Work." *CSCW '06 Proceedings of the 2006 20th anniversary conference on Computer supported cooperative work,* 515-524.

Yasouka, Mika, and Pernille Bjorn. 2011. "Machine Translation Effects on Communication: What Makes it Difficult to Communicate through Machine Translation?" *Second International Conference on Culture and Computing*: 110-115. Kyoto, Japan.

PUBLICATION V

Machine translation and fair access to information

Mary Nurminen and Maarit Koponen

"Towards ethical, sustainable machine translation." Special issue of Translation Spaces 9(1): 150–169
https://www.benjamins.com/catalog/ts

Nurminen, Mary and Maarit Koponen. (2020) Machine Translation and Fair Access to Information. *Translation Spaces* 9(1): 150–169. doi: https://doi.org/10.1075/ts.00025.nur

Machine Translation and Fair Access to Information

Mary Nurminen* and Maarit Koponen**

*Tampere University
**University of Turku and University of Helsinki

Abstract

This article contributes to the discussion on fairness and ethics in MT by highlighting efforts that have been made to use MT for the humanitarian purpose of increasing accessibility to information for groups that are underserved. The article provides an overview of example projects in which MT has been implemented for this purpose in three contexts: civic participation, public health and safety, and media and culture. In addition, the article examines some of the ethical issues surrounding efforts to use MT for accessibility, including issues of quality, acceptability, and the need to involve stakeholders in development.

Keywords

machine translation, accessibility, ethics, machine translation for information purposes, post-editing, linguistic accessibility

A number of motivations have fueled the research and development of machine translation (MT), including those from science, defense, government, and industry (Melby and Warner 1995; Hutchins 2010). A further motivation has been the idea that MT could be used to benefit humankind by promoting accessibility to information. Especially in the early days of development, mentions were made of researchers being motivated by "idealism: the promotion of international cooperation and peace, the removal of language barriers, the transmission of technical, agricultural and medical

information to the poor and developing countries of the world." (Hutchins 1986, 15) However, research interests quickly became devoted to methods, technologies and evaluations of different systems, with little attention given to utopian motivations. In the late 1990s, research emerged on the need for language technology solutions for minority languages and immigrant populations, and the humanitarian motivations for MT resurfaced. Coinciding with this was the rise of artificial intelligence systems and a subsequent re-energization of discussions on ethical issues. All of these actions have led to the current examination of ethics and fairness in machine translation.

As part of this examination, we want to reconsider the utopian idea that, in addition to providing commercial enterprises with a way to increase efficiency to reduce translation costs, MT can be used for humanitarian purposes. The IEEE Global Initiative for Ethical Consideration of Artificial Intelligence and Autonomous Systems (A/IS) lists three general principles related to the use of artificial and autonomous systems, stating that they should: "1) embody the highest ideals of human beneficence as a superset of Human Rights, 2) prioritize benefits to humanity and the natural environment from the use of A/IS, and 3) mitigate risks and negative impacts, including misuse, as A/IS evolve as socio-technical systems" (IEEE 2017, 23). The first goal of this paper is to bring into the discussion on ethics and MT part of the second IEEE general principle, namely, the benefits to humanity that are enabled by MT. Access to information can be seen as a human right necessary for participation in society and as a means of ensuring equality, and efforts to improve accessibility to information can thus be viewed as beneficial to society as a whole. Various governmental institutions, non-governmental organizations, private businesses and academic researchers have indeed stepped in to develop tools to address the issue of multilingual access to information and help ensure accessibility for various, previously underserved, groups. A second goal of the paper is to introduce some of the ethical issues that might arise in efforts to use MT for accessibility purposes. We do not aim to provide a comprehensive account of all past and ongoing projects, nor all ethical issues. Rather, our objective is to raise awareness of the concrete actions being taken to use MT to benefit humankind, and to introduce some of the ethical issues that may be involved in these actions.

Because many of the projects we introduce in this paper are recent or ongoing, our data consisted of websites, white papers, project proposals, reports and personal communications in addition to scientific articles. The structure of the paper is as

follows: the first section describes the phenomenon in question and the main needs it sets out to meet. Section 2 introduces example projects that use MT to improve accessibility to civic participation, health and safety information, and media and culture. In Section 3 we review some ethical issues that may arise in implementing MT for accessibility, and Section 4 contains our conclusions and recommendations.

1. Accessibility to information and machine translation

In this paper we examine efforts that aim to increase target audiences' accessibility to the information they need by lowering or removing language barriers through the use of MT. Barriers to accessibility can be defined in many ways, and various recommendations and requirements have been drafted to remove them (see, for example, European Commission 2015). While measures aiming to remove barriers to accessibility often focus on sensory and physical or cognitive barriers, another factor is language and linguistic accessibility (see Matamala and Ortiz-Boix 2016, 13–14), and this factor is the focus of this paper. The audiences targeted in linguistic accessibility initiatives include groups who need certain types of information but are not able to access or read that information because of insufficient command of the languages it is offered in. Such groups include, for example, refugees, migrants, and people in crisis situations.

1.1 Access to information as a human right and language as an obstacle to accessibility

The right to information is both a practical question and one of the principal human rights that can be derived from international agreements such as the UN's Universal Declaration on Human Rights (United Nations 1948) and the European Convention on Human Rights (Council of Europe 1950). Although the European Convention on Human Rights, for example, does not list the right to obtain information as such, in the case law of the European Court of Human Rights this right has been upheld on the basis of other articles of the human rights convention, particularly Article 10, which guarantees freedom of expression, but also Article 2 (right to life), Article 6 (right to a fair trial), and Article 8 (right to private and family life) (Tiilikka 2013). Tiilikka (2013, 102) argues that public authorities can be considered to have an obligation to provide information, and the promotion of access to information is often covered not only in international agreements, but also in national laws.

In part, the obligation to provide information concerns openness and availability of official information, but availability alone does not guarantee accessibility. A key point for linguistic accessibility is the language (or languages) in which information is available. When information is provided only in certain languages, insufficient knowledge of those languages may become a barrier to accessibility. The role of multilingualism as essential for disseminating information and knowledge is widely acknowledged. For example, the European Union considers supporting multilingualism to be important for promoting both cultural identity and social cohesion (European Commission 2012, 2). Accessibility of information in multiple languages, and multilingual practices as a whole, can be seen as "a basic condition for the development of a truly inclusive knowledge society" (UNESCO Executive Board 2007, 2).

1.2 The challenges of multilingualism and MT as a potential solution

In recent years, growing global movement has increased the need for multilingual information, both in terms of the amount of text and the variety of languages. Translation and interpreting are increasingly needed during immigration and refugee processes, and on a continuing basis to make social services and education accessible to immigrants (Biel and Sosoni 2017, 354). The provision of multilingual information is, however, limited by the resources required. One argument raised against multilingualism has been the costs of producing information in multiple languages through translation and other measures, and arguments have been made for using only official language(s) or a lingua franca, often English. However, reliance on English as a "common language" is not unproblematic. For example, a survey of European Union residents found that, of those who did not speak English as a native language, only 38% indicated they were able to hold a conversation in English, and only 25% were able to read a newspaper or follow news on television or radio (European Commission 2012, 21, 28–29). In a study carried out among the refugee and migrant population in Greece, 65% did not understand spoken English, and 80% did not understand written English (Ghandour-Demiri 2017, 16–17).

As a counterpoint to the costs of multilingualism, Gazzola and Grin (2013, 99) note that focusing only on the monetary costs related to translation and interpreting ignores other factors that are relevant when comparing multi- and monolingualism. Factors related to monolingualism include: the potential costs for the (governmental or other)

organization caused by misunderstandings, exhaustion and lower efficiency when workers at that organization who speak a language other than the official language are forced to constantly operate in the official language; costs incurred by the people who need to communicate with that organization in a language they do not know; and psychological costs due to exclusion (Gazzola and Grin 2013, 99–100).

Technology, particularly MT, could offer one solution for the problem of limited human and monetary resources. However, a long-recognized problem in the area of MT for accessibility is a discrepancy between the languages needed in accessibility solutions, for example, the languages spoken in countries from which large numbers of people are currently migrating, and the languages for which solutions are being developed for commercial or large organizational use (Somers 1997). As Carbonell et al. (2006, 120) pointed out, "[u]nfortunately, such economic imperatives exclude most minority languages where MT is most needed for humanitarian purposes." Somers (1997, 11) predicted that only governmental agencies will fund solutions for these languages "unless the private sector sees this as an area where it can make charitable donations." Over the past decade, attention has turned to the question of language technology solutions and resources for languages that have been underserved. For example, in 2017 the non-profit organization Translators without Borders launched the Gamayun initiative, the focus of which is to "bring language technology to bear for the world's most marginalized communities" (Ansari and Petras 2018, 2). Interest in these languages has also risen in MT research, as is evidenced by two recent conference workshops devoted to MT for "low resource languages."[1,2]

2. Example projects using MT to promote fair access to information

Over the years MT has been a component in a number of concrete projects aimed at increasing accessibility to information. Some of these projects have a broad scope, with a variety of target groups, accessibility issues to be addressed, and tools involved. Others focus specifically on MT, as a productivity boost for translators' work or as a tool used directly by target groups to access information. In our research, we aimed to include projects that involved not only the development of MT solutions,

[1] https://sites.google.com/view/loresmt-2018/

[2] https://sites.google.com/view/loresmt/

but also well-defined use cases and target users for those solutions. This section introduces several example projects which aim at increasing accessibility to civic participation, health and safety information, and culture and media.

2.1 Accessibility to civic participation

Projects implementing MT to promote accessibility to civic participation have been undertaken by large governments as well as smaller national bodies. One of the pioneers in this area is the European Union, where efforts to develop MT started as early as 1976 in the European Commission (European Commission 2010, 68). As the EU grew, with more member states, languages, and information to translate, the development of machine translation continued and its use accelerated. The current iteration of MT technology, eTranslation, relies on neural MT and is used by translators as a part of the translation workflow, with material being first machine translated and then post-edited before publication. eTranslation is also offered as a service to people working in public administrations in the EU, Norway and Iceland[3]. Civil servants can access eTranslation directly or it can be integrated with other EU information systems. As with many activities of the European Union, the ultimate goals of the eTranslation service concern multilingual inclusion and participation. Other very large multinational organizations such as the United Nations have also implemented MT systems in their translation workflows (Pouliquen et al. 2013). Projects on a smaller scale also have similar goals of participation and civic engagement, aiming to deliver more information to users in official languages or to ensure access to information for new, previously underserved, groups of people, including speakers of non-official languages, minority languages, and minority languages that arrive with migrant populations.

In the early 2010s, Latvia was facing the need to provide information for a large Russian-speaking minority and also to provide government materials in English to facilitate the country's growing role in the European Union. To address both of these issues, the government decided to implement MT "to facilitate communication between the Latvian public sector and Latvia's citizens" (Vasiljevs et al. 2014, 183). Working with a language technology provider, an MT system was developed to translate between Latvian and Russian and between Latvian and English. Public

[3] https://ec.europa.eu/cefdigital/wiki/display/CEFDIGITAL/Machine+translation

administration employees can use the tool to translate official texts, which they then post-edit for publication. For the general public, the tool also provides an internet portal where anyone can enter text, documents or websites to be machine translated for gisting purposes.

Another project born out of a need for improved accessibility to public services, for recent immigrants as well as people with reduced reading capabilities, was the DigInclude project in Sweden, started in 2016. This cooperative effort between university researchers, language technology providers, and governmental organizations, aimed to improve civic engagement through the development of linguistic resources and tools for writing support, text simplification, multilingual terminology, and MT (Jönsson 2016). At the time of writing, two phases of the project had been completed, but funding for a third had not been secured. However, the MT part of the project has continued in two follow-up projects between the MT provider and the Swedish Migration Agency. The goal in these new projects is to develop MT solutions between Swedish, English and Arabic, to be used initially in translating the agency's website and, in the second project, for translating documents related to applications for asylum and citizenship[4].

The MuTUAL project addressed Japanese municipalities' need to offer official information in languages other than Japanese, which was deemed challenging to accomplish through human translation or post-editing due to expense and time constraints (Miyata, Hartley, Paris, et al. 2015). The MuTUAL system aimed to create tools that would help content creators to produce public information in a form that would optimize MT output quality, enabling them to distribute information in unedited MT form (Miyata, Hartley, Kageura, et al. 2016). Currently, no municipality has implemented the MuTUAL system although many offer access to MT tools integrated into their websites. There are, however, plans to implement and evaluate the MuTUAL system in real-world situations in the future[5].

2.2 Accessibility to health and safety information

Difficulty in accessing health information, resources and services due to language barriers has been linked with negative consequences, such as receiving less

---

[4] Anna Sågvall Hein, email message to authors, October 8, 2019

[5] Rei Miyata, email message to authors, October 3, 2019

preventative healthcare (see Dew et al. 2018, 57). Some research has been conducted on the use of MT for increasing accessibility to health information, although in their recent review, Dew et al. (2018, 60) found that most publications on this topic involve pilot studies or evaluations of the feasibility of existing tools, often freely available online systems like Google Translate (e.g. Das et al. 2019). Work has focused on text translation, although some projects, such as BabelDr (Bouillon et al. 2017), have involved automatic speech translation systems to assist in communication between doctors and patients.

One of the pioneers in implementing MT, the Pan American Health Organization (PAHO) aimed specifically at increasing the spread of health information throughout North and South America. MT development in the organization began in 1976 and by 1980 MT and post-editing were part of the translation process (Vasconcellos and León 1985). By 2009, MT was fully integrated into the translation workflow and used to process more than 90% of translation jobs (Aymerich and Camelo 2009). Another early international MT and public health initiative was the Global Public Health Intelligence Network GPHIN (Blench 2008), developed by the Public Health Agency of Canada in partnership with the World Health Organization. This system, still in use today, monitors global media sources and websites to identify reports of disease outbreaks and other potential public health events, utilizing MT to translate English to/from Arabic, Chinese, Farsi, French, Russian, Portuguese and Spanish (Blench 2008).

The TransPhorm project[6] (Kirchhoff et al. 2011), carried out in the US, involved a collaboration between university researchers and health departments at the state and local level and investigated the use of MT in the public healthcare sector, focusing on MT from English into Spanish and Chinese. The work carried out in the project involved testing the feasibility of generic and domain-specific MT tools and post-editing for disseminating public health information. The work included MT and post-editing experiments for translating English content into Spanish for public health information websites as well as for health promotion materials used by public health departments (Turner, Bergman, et al. 2014). A web-based translation system was developed in which public healthcare workers could produce multilingual material by

[6] http://www.nwcphp.org/research/projects/current/transphorm-machine-translation-of-public-health-information

using MT and post-editing (Laurenzi et al. 2013). Turner, Choi, et al. (2019) further evaluated Google Translate and a domain-specific system for communication between emergency medical services and patients in Chinese and Spanish.

The recent EU-funded Health In My Language project[7] developed MT systems adapted for translating public health information from English into Czech, German, Polish and Romanian. The project involved collaboration between academia and two partners: a website that publishes systematic reviews of medical studies[8], and regional public health service providers in Scotland. One use case involved professional translators post-editing MT output for plain language summaries of medical reviews to be published on the medical review website, while other use cases investigated the feasibility of publishing information in the form of unedited MT either on the medical review website or on the website of the regional health services (Birch, Ried, et al. 2018).

MT has also been implemented to improve access to safety information and resources. It has proven to be helpful in ensuring accessibility in circumstances in which information plays a critical role, specifically in crisis situations. Since 2014, the Words of Relief program[9] of Translators without Borders has been instrumental in mediating information in a number of natural disasters, public health and refugee situations through translation and interpreting services and training. The program uses language technology to enable quicker responses to information needs. Their Kató translation platform[10] includes MT alongside a variety of other tools used by their community. The EU-funded INTERACT (International Network on Crisis Translation) project[11] was initiated in April 2017 to specifically address translation needs in crisis scenarios. The project is being conducted in cooperation between partners from universities, private companies, and NGOs, and includes the development of MT systems and processes for use in crisis scenarios. Crisis translation and MT in that context have recently attracted growing research interest and is emerging as a field of its own (see, for example, Federici and O'Brien 2020).

---

[7] http://www.himl.eu/

[8] http://www.cochrane.org

[9] https://translatorswithoutborders.org/our-work/crisis-response/

[10] https://translatorswithoutborders.org/our-work/kato-translation-platform/

[11] https://sites.google.com/view/crisistranslation/home

2.3 Accessibility to culture and media

Accessibility to media, including audiovisual (AV) media like television or films, is another factor recognized as important for social inclusion and for promoting intercultural dialogue (see Matamala and Ortiz-Boix 2016). Accessibility of news content is also very relevant for participation in society. Audiovisual media poses its own set of challenges, as obstacles to accessibility may be sensorial (hearing, vision) in addition to linguistic. Various projects have aimed to promote linguistic accessibility of AV media through the use of MT.

Some projects have explored fully automated generation and translation of subtitles, such as the MUSA project[12] where speech recognition and MT were used to create English, French and Greek subtitles for documentaries and current affairs programs (Piperidis et al. 2004). The EU-funded HBB4All project[13] also investigated English-to-Spanish MT as well as automatic intralingual English subtitling through automatic speech recognition as a way to increase access to news content originally produced in English (Matamala, Oliver, et al. 2015).

Other projects have focused on the use of MT post-editing as a way to increase productivity. The eTITLE project (Melero et al. 2006), a collaboration between university researchers and broadcasting companies in Spain and the Czech Republic, developed a web-based subtitling platform where MT was offered as a tool for subtitlers alongside translation memories for language pairs involving English, Spanish, Catalan and Czech. MT post-editing of subtitles was also investigated in the EU-funded SUMAT project[14], conducted by academia and industry partners in the fields of subtitling and media/video content production. The project developed MT systems for subtitling in language pairs including English, German, French, Spanish, Swedish, Portuguese, Dutch, Serbian and Slovenian, and included a large-scale evaluation where professional subtitlers tested the usability of MT and post-editing for subtitling (Bywood, Georgakopoulou and Etchegoyhen 2017, 496–497).

While most projects aiming to address accessibility of audiovisual media have focused on MT for subtitles, the Spanish ALST Linguistic and Sensorial Accessibility project investigated the use of MT and post-editing for voice-overs (Ortiz-Boix and

[12] http://sifnos.ilsp.gr/musa

[13] http://pagines.uab.cat/hbb4all/

[14] http://www.fp7-sumat-project.eu/

Matamala 2017) and audio description (Matamala and Ortiz-Boix 2016) from English to Spanish and from Catalan to Spanish. Audio description refers to a process where "an oral explanation of the most relevant visuals (characters, settings, actions, etc.)" is inserted into audiovisual products such as films, making them accessible to blind and visually impaired audiences (Matamala and Ortiz-Boix 2016, 13). Based on their preliminary experiments, Matamala and Ortiz-Boix (2016, 22) suggest that MT in combination with audio description offers a way to increase multilingual accessibility of AV content.

The EU-funded MeMAD[15] project, which involves partners from academia, public service broadcasting and audiovisual archiving in Finland and France, as well as language service and technology companies, aims to increase the accessibility of audiovisual and audio material through automatic speech recognition, automatic video content description and machine translation. The project investigates the use of MT for subtitling, but also for other content like audio description and (human or machine generated) content descriptions of videos. For subtitling, the use cases defined involve both the use of MT and post-editing by professional subtitlers working for public service broadcasters and the use of fully automatic translation of subtitles as potential ways to increase accessibility to news, current affairs, and cultural programming for minority language speakers (Braeckman et al. 2019, 29–32). The language pairs for MT include Finnish, Swedish, English, French and Dutch (Braeckman et al. 2019, 48).

Another recent EU project, GoURMET[16], also brings together academia and public service broadcasters from Germany and the UK, and focuses on MT tools that would enable the production of content in selected low-resource languages as well as the monitoring of media, such as business news, in these languages. The language pairs covered include English to/from Turkish, Gujarathi, Swahili and Bulgarian (Birch, Haddow, et al. 2019).

[15] https://memad.eu/

[16] https://gourmet-project.eu/

3. Ethical considerations in MT for accessibility

As mentioned earlier, besides providing the case of MT for accessibility as part of the higher-level discussion on MT and ethics, a second goal of this paper is to explore ethical issues involved in using MT for accessibility purposes. We identify here potential ethical issues related to quality, acceptability, and the inclusion of key stakeholders. However, this list is meant as a starting point for discussion and not an exhaustive list of issues.

3.1 Quality issues

A common method of using MT for accessibility is to have raw MT output post-edited by humans before being disseminated to end users. This process is increasingly used in the translation industry and has been found to increase productivity (see e.g. Plitt and Masselot 2010, 10), suggesting that MT with post-editing could also promote accessibility by making it possible to translate more content with the same resources. Post-editing can take different forms depending on issues such as the context in which the translation will be read and the level of quality that is required. A common distinction is made between "full" (or "maximal") post-editing, which entails more thorough editing aiming for "publishable quality," and "light" (or "rapid") post-editing, which aims to ensure that information is accurate and comprehensible, but does not include stylistic improvements to the MT output (for an overview, see Hu and Cadwell 2016). However, the extent to which post-editing can improve productivity depends on MT quality, which varies for different language pairs and text types. Furthermore, productivity gains are affected by conditions such as whether experienced professional post-editors are involved and whether the MT system is tailored for the content type in question (García 2011, 228).

Many of the projects covered in Section 2 involved MT and post-editing. In the Health In My Language project, post-editing MT output of medical summaries was shown to increase productivity and received favorable feedback from professional translators in all language pairs except English-to-Polish, where MT quality was considered too poor (Birch et al. 2018, 16). Turner, Bergman, et al. (2014) also suggest that based on the TransPhorm project, MT and post-editing by bilingual public health staff—rather than professional translators—could be a feasible method for producing multilingual health promotion materials. Although the use of MT and post-editing has been less common for audiovisual translation than for more

conventional text, it was seen as a viable solution for subtitling in the SUMAT project evaluations (Bywood, Georgakopoulou and Etchegoyhen 2017, 504), and the ALST project suggests its feasibility for voice overs and audio description also (Matamala and Ortiz-Boix 2016; Ortiz-Boix and Matamala 2017).

Another way of using MT is to distribute information in its raw, unedited form. In these cases, quality is a critical question in two ways. First, information in the form of poor MT output may be very difficult to read or even unintelligible, thereby forming an obstacle to accessibility. Second, unedited MT might contain errors that result in misinformation. Even with the improvements in quality reported recently, particularly for neural MT approaches, MT output remains imperfect. Recent studies suggest that the adoption of neural MT approaches can produce more fluent and idiomatic output compared to phrase-based statistical models. Castilho et al. (2017, 118), however, point out that the improved fluency of neural MT is not necessarily accompanied by a corresponding improvement in adequacy. For example, neural MT output may involve more frequent omissions of words or even longer passages (Castilho et al. 2017). Such omissions can lead to missing information, which is particularly difficult to recover without access to or understanding of the source text (Koponen and Salmi 2015).

Furthermore, more fluent translations can in fact make errors of meaning more difficult to detect. This in turn exacerbates the risk of misinformation, as the apparently fluent output may foster unwarranted trust in the MT. In their study of MT and trust, Martindale and Carpuat (2018) indeed observed that participants reacted much more strongly to errors in fluency than errors in meaning, pointing out that the participants were perhaps not even aware of some of the content errors. Even if apparently fluent, and therefore convincing, incorrect translations are relatively rare, they can still be particularly dangerous as they mislead the reader (see Martindale et al. 2019). Although some recent studies (e.g. Martindale and Carpuat 2018) have looked into reading comprehension with neural MT, the effect of such potentially misleading errors on a reader using raw machine translated content for information purposes is largely unexplored. Particularly in situations where a misunderstanding could lead to harm for the person accessing the information, or for others, raw MT alone is not sufficient.

3.2 Acceptability of the use of MT

Two important issues of acceptability need to be considered when using MT for increasing accessibility to information. The first involves data privacy. As is well established, while free online MT systems are convenient, there are privacy issues and potential data breach issues involved in their use. Care must be taken to ensure that systems developed for accessibility purposes offer the levels of data privacy and protection required, with much depending on the types of information involved. For example, data privacy requirements may be lower for general information and instructions concerning immigration, whereas content relating to specific people involves a need for very high protection. For these reasons, many of the projects introduced in Section 2 involve protected or proprietary MT systems.

The second issue of acceptability we see concerns how acceptable the practice is to the target audience. Their perception of acceptability can be affected by several factors, including the quality of the MT output, the situations in which MT is used, the text types involved, and the purpose users have for certain texts in their own language.

This perspective was addressed most extensively in the context of civic participation in two studies carried out in Canada, investigating the feasibility of using raw and post-edited MT to increase the accessibility of information. Bowker (2009) addressed information provided online by municipal or provincial governments to "official language minority communities" of Canada (French-speaking communities in predominantly English-speaking areas and vice versa), whereas Bowker and Buitrago Ciro (2015) addressed information offered by a public library to newly arrived Spanish-speaking immigrants. Both studies focused on a recipient evaluation, in which participants from the target communities read translated versions as raw MT, rapidly post-edited MT, maximally post-edited MT and human translation, and assessed which version met their needs best, also considering the time and cost needed to produce each version (Bowker 2009, 142; Bowker and Buitrago Ciro 2015, 179). The findings of the two studies differed. In the study involving official language minority communities, neither group considered raw MT acceptable. The French-speaking participants showed a clear preference for human translation, whereas nearly half of the English-speaking participants considered the rapidly post-edited MT suitable for their needs (Bowker 2009, 142, 146). In the later library study, the majority indicated that rapidly post-edited MT met their needs best, and depending on

text type, up to 38% even considered raw MT most suitable (Bowker and Buitrago Ciro 2015, 180). These findings suggest that post-edited and even raw MT can be a viable solution for increasing access to information. Users of the information may even find these solutions preferable when considering the time and cost aspects, which affect how fast the information can be updated and how much of it can be translated (Bowker and Buitrago Ciro 2015, 179).

The lower acceptance observed in Bowker (2009) may reflect differences in the quality of the MT, which likely improved between the older and newer study due to developments in MT models. However, Bowker (2009) also proposed that acceptability is also influenced by the reasons people have for wanting to have information in their own language: if the purpose is to vindicate the rights and preserve the culture of a recognized language minority, users want all official information as well-formed human translations. This also relates to an important question of perception. Having information available in "only" machine translated form may lead to a perception of that language as less important than languages for which the same information is translated by humans. Close attention, therefore, needs to be paid to the needs and perspectives of the people targeted by information accessibility efforts.

The needs of users accessing public health information and the acceptability of raw MT were also investigated in the Health In My Language project. Again, the reasons for accessing information and the context where it was offered appeared to affect the results. In a survey conducted with people using the medical review website to read plain language summaries, most users (75% of German users and nearly 50% of Czech and Romanian users) found raw MT output acceptable, but due to lower quality, raw MT was acceptable to only 6% of Polish users (Birch et al. 2018, 20). On the other hand, on the website of the National Health Services in Scotland, Romanian and Polish users found the use of raw MT unacceptable and indicated that they would expect only fully accurate health advice to be provided by the public health services (Birch et al. 2018, 27). Overall, information in the medical setting is particularly critical. For example, Kirchhoff et al. (2011) argue that the use of unedited MT for health information is never acceptable, so MT output must be always post-edited. In a more recent experiment involving "anticipatory guidance" resources for child well-being, Das et al. (2019, 249) also found unedited MT to contain inaccuracies that could "pose significant risks to child health outcomes" and therefore argued for the

need for high-quality human translations. Similar findings were reported by Turner, Choi, et al. (2019, 11), who found MT too risky for communication between emergency medical services and patients due to potentially critical errors.

3.3 Involvement of stakeholders in development efforts

A common thread seen in research and projects to implement MT for accessibility is that resourcing and financial support are provided by a consortium of stakeholders from different areas such as governments, NGOs, universities and the private sector. Contrary to earlier views that solutions for humanitarian purposes might be supported only by government agencies (see Carbonell et al. 2006; Somers 1997), most of the projects described in Section 2 involved collaboration between various of these different stakeholders. Also, an analysis of the 19 papers accepted for the two workshops on MT for low resource languages mentioned in Section 1.2 reveals not only a variety of languages being studied, but also a variety of funding sources for research efforts. In addition to the universities that traditionally provide human resources and funding for research, these projects were supported by private companies, national governments, the European Union, and non-profit organizations for the advancement of specific languages (Liu 2018; Liu and Karakanta 2019). Possible explanations for this tendency toward collaboration might be a recognition of the need for diverse competences to achieve the goals of such initiatives, the need to pool resources, or stakeholders' desire to have an influence on the outcomes of the projects. It also reflects the early motivations to enlist technologies not only for commercial gain but also for humanitarian purposes.

Other stakeholders in MT for accessibility projects are the eventual users of the systems to be developed and the eventual readers of the information to be produced, and the inclusion of these groups has been called for by some of these projects. For example, the INTERACT project's Ethics Recommendations for Crisis Translation Settings emphasize the importance of considering the needs of the affected communities when developing translation technology for crisis situations, which "requires consultation with and training of users as well as community-based evaluation of such technologies through participative research practices." (O'Mathúna et al. 2019, 8) A community-based participatory research framework was also employed by Bowker and Buitrago Ciro (2015) in order to best meet the needs of

potential future users, and a survey of user needs was conducted in the Health In My Language project (Birch et al. 2018).

Recommendations have also been made to involve language professionals in projects. For example, Parra Escartín and Moniz (2020), addressing the context of crisis translation, note the value of professional translators in curating and managing data quality. Current MT technology also relies on large amounts of training data in the form of translated texts (parallel corpora). As pointed out by Parra Escartín and Moniz (2020, 135), ethical considerations related to data, such as ownership of data and acknowledgement of the contribution of the translators who created it, should not be ignored. Furthermore, they highlight the need to curate such training data, particularly the need to remove/anonymize potentially sensitive information.

4. Conclusions

In this article we contributed to the ongoing discussion of ethics and MT by highlighting example efforts to use MT for humanitarian purposes, specifically for increasing accessibility to information for groups that were previously underserved, and by examining ethical issues in efforts to use MT for accessibility. We reviewed a number of projects implementing MT for increasing accessibility to civic participation, health and safety information, and culture and media, and then discussed the ethical issues of quality, acceptability and stakeholder participation.

One limitation of our work is that it is not based on a systematic review of projects. Mostly because of the recent or ongoing status of many projects of this sort, we relied on convenience sampling, personal communications, and other less formal methods for data gathering. In the future, a more systematic approach could be employed to compile a comprehensive list of projects that use MT for accessibility purposes. Another area for future development is the contemplation of ethical guidelines for using MT to improve accessibility. A good starting point might be the INTERACT project's ethics recommendations, specifically those related to translation technology development and deployment (O'Mathúna et al. 2019).

Based on our analysis of the MT projects discussed in this article, we propose four conclusions for future consideration. First, projects have benefited from the inclusion of a variety of stakeholders from different areas. Involving the organizations responsible for promoting accessibility, groups with technical expertise, future users of systems, and language professionals is a best practice. These best practices should

also address ethical factors such as protection of potentially sensitive data and proper acknowledgement of stakeholder contributions. Second, although accessibility efforts in individual projects may involve low-resource languages that are unfamiliar to the project team, there may be other projects working with the same languages which also have limited resources. Collaboration between organizations could be mutually beneficial. Third, we recognize that there are contexts in which raw, unedited MT might be useful. However, we encourage stakeholders to carefully analyze contexts for suitability before implementing raw MT, taking into consideration quality issues and other aspects of acceptability.

A final observation is that finding information on projects that were in the planning or implementation phases proved much easier than finding follow-up information on the eventual use and users of the systems produced in those projects. It was not always clear whether this was due to projects being discontinued or if research on use and users was limited. In the future, we would like to see research focus continue throughout the implementation of systems and we would also like to see more research on the users of those systems.

Acknowledgements

This work has been partially performed while Maarit Koponen was working for the project "Methods for Managing Audiovisual Data," supported by the European Union's Horizon 2020 Research and Innovation Programme under Grant Agreement No 780069.

References

Ansari, Aimee, and Rebecca Petras. 2018. "Gamayun: The Language Equality Initiative." Translators Without Borders. Accessed 15 October 2019, https://translatorswithoutborders.org/wp-content/uploads/2018/03/Gamayun-Language-Equality-Initiative-March-2018.pdf.

Aymerich, Julia, and Hermes Camelo. 2009. "The Machine Translation Maturity Model at PAHO." In *Proceedings of the Twelfth Machine Translation Summit*, 403–409.

Biel, Łucja, and Vilelmini Sosoni. 2017. “The Translation of Economics and the Economics of Translation.” *Perspectives* 25 (3): 351–361. DOI 10.1080/0907676X.2017.1313281.

Birch, Alexandra, Barry Haddow, Ivan Tito, Antonio Valerio Miceli Barone, Rachel Bawden, Felipe Sánchez-Martínez, Mikel L. Forcada, Miguel Esplà-Gomis, Víctor M. Sánchez-Cartagena, Juan Antonio Pérez-Ortiz, Wilker Aziz, Andrew Secker, and Peggy van der Kreeft. 2019. “Global Under-Resourced Media Translation (GoURMET).” In *Proceedings of Machine Translation Summit XVII Volume 2: Translator, Project and User Tracks*, 122. European Association for Machine Translation.

Birch, Alexandra, Juliane Ried, Colin Davenport, Matthias Huck, and David Mareček. 2018. “D5.6: Report on Third Year’s Evaluation.” Health In My Language Project. http://www.himl.eu/files/D5.6_Final_Evaluation.pdf.

Blench, Michael. 2008. “Global Public Health Intelligence Network (GPHIN).” In *Eighth Conference of the Association for Machine Translation in the Americas*.

Bouillon, Pierrette, Johanna Gerlach, Hervé Spechbach, Nikos Tsourakis, and Sonia Halimi. 2017. “BabelDr vs Google Translate: A User Study at Geneva University Hospitals (HUG).” In *The 20th Annual Conference of the European Association for Machine Translation: User Studies and Project/Product Descriptions*, 47–52.

Bowker, Lynne. 2009. “Can Machine Translation Meet the Needs of Official Language Minority Communities in Canada? A Recipient Evaluation.” Linguistica Antverpiensia New Series-Themes in *Translation Studies* 8: 123–155.

Bowker, Lynne, and Jairo Buitrago Ciro. 2015. “Investigating the Usefulness of Machine Translation for Newcomers at the Public Library.” *Translation and Interpreting Studies* 10 (2): 165–186. DOI 10.1075/tis.10.2.01bow.

Braeckman, Karel, Simon Debacq, Harri Kiiskinen, Nico Oorts, Lauri Saarikoski, Raphaël Troncy, Wim Van Lancker, Dieter Van Rijsselbergen, Maarten Verwaest, and Kim Viljanen. 2019. “D6.4 Specification of the Data Interchange Format, Intermediate Version.” MeMAD Project. https://memad.eu/wp-content/uploads/D6.4-Specification-of-the-data-interchange-format-intermediate-version-version-1.0.pdf.

Bywood, Lindsay, Panayota Georgakopoulou, and Thierry Etchegoyhen. 2017. "Embracing the Threat: Machine Translation as a Solution for Subtitling." *Perspectives: Studies in Translatology* 25 (3): 492–508. DOI 10.1080/0907676X.2017.1291695.

Carbonell, Jaime G., Alon Lavie, Lori Levin, and Alan W. Black. 2006. "Language Technologies for Humanitarian Aid." In *Technology for Humanitarian Action*, edited by Kevin M. Cahill, 111–138. New York, USA: Fordham University Press & The Center for International Health and Cooperation.

Castilho, Sheila, Joss Moorkens, Federico Gaspari, Iacer Calixto, John Tinsley, and Andy Way. 2017. "Is Neural Machine Translation the New State of the Art?" *The Prague Bulletin of Mathematical Linguistics* 108 (1): 109–120. DOI 10.1515/pralin-2017-0013.

Council of Europe. 1950. "European Convention on Human Rights (European Convention for the Protection of Human Rights and Fundamental Freedoms, as Amended by Protocols Nos. 11 and 14)." Accessed 20 February 2020, https://www.echr.coe.int/Documents/Convention_ENG.pdf.

Das, Prithwijit, Anna Kuznetsova, Meng'ou Zhu, and Ruth Milanaik. 2019. "Dangers of Machine Translation: The Need for Professionally Translated Anticipatory Guidance Resources for Limited English Proficiency Caregivers." Clinical Pediatrics 58 (2): 247–249. DOI 10.1177/0009922818809494.

Dew, Kristin N., Anne M. Turner, Yong K. Choi, Alyssa Bosold, and Katrin Kirchhoff. 2018. "Development of Machine Translation Technology for Assisting Health Communication: A Systematic Review." *Journal of Biomedical Informatics* 85 (July): 56–67. DOI 10.1016/j.jbi.2018.07.018.

European Commission. 2010. "Translation at the European Commission – a History." Luxembourg. DOI 10.2782/16417.

European Commission. 2012. "Special Eurobarometer 386: Europeans and Their Languages." Accessed 20 February 2020, https://ec.europa.eu/public_opinion/archives/ebs/ebs_386_en.pdf.

European Commission. 2015. “Directive of the European Parliament and the Council on the approximation of the laws, regulations and administrative provisions of the Member States as regards the accessibility requirements for products and services.” Accessed 20 February 2020, https://eur-lex.europa.eu/eli/dir/2019/882/oj.

Federici, Federico M. and Sharon O'Brien. 2020. *Translation in Cascading Crises*. London: Routledge.

García, Ignacio. 2011. “Translating by Post-Editing: Is It the Way Forward?” *Machine Translation* 25 (3): 217–237. DOI 10.1007/s10590-011-9115-8.

Gazzola, Michele, and François Grin. 2013. “Is ELF More Effective and Fair than Translation? An Evaluation of the EU’s Multilingual Regime.” *International Journal of Applied Linguistics* 23 (1): 93–107. DOI 10.1111/ijal.12014.

Ghandour-Demiri, Nada. 2017. “Language & Comprehension Barriers in Greece’s Migration Crisis: A Study on the Multitude of Languages and Comprehension of Material Provided to Refugees and Migrants in Greece.” Translators without Borders and Save the Children. Accessed 1st March 2020, https://resourcecentre.savethechildren.net/node/12449/pdf/language-comprehension-barriers.pdf.

Hu, Ke, and Patrick Cadwell. 2016. “A Comparative Study of Post-Editing Guidelines.” In *Proceedings of the 19th Annual Conference of the European Association for Machine Translation*, EAMT 2016, 346–53.

Hutchins, William John. 1986. *Machine Translation: Past, Present, Future*. Chichester: Ellis Horwood.

Hutchins, William John. 2010. “Machine Translation: A Concise History.” *Journal of Translation Studies. Special Issue: The Teaching of Computer-Aided Translation*. 13 (1–2): 29–70.

IEEE. 2017. “Ethically Aligned Design: A Vision for Prioritizing Human Well-Being with Autonomous and Intelligent Systems.” *The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems*. https://ethicsinaction.ieee.org/#read

Jönsson, Arne. 2016. “DigInclude – Digital Inclusion in the Networked Society for Groups with Special Needs (Proposal for Vinnova).” Reference accessed 1st March 2020, https://liu.se/en/research/diginclude.

Kirchhoff, Katrin, Anne M. Turner, Amittai Axelrod, and Francisco Saavedra. 2011. “Application of Statistical Machine Translation to Public Health Information: A Feasibility Study.” *Journal of the American Medical Informatics Association* 18 (4): 473–478. DOI 10.1136/amiajnl-2011-000176.

Koponen, Maarit, and Leena Salmi. 2015. “On the Correctness of Machine Translation: A Machine Translation Post-Editing Task.” *Journal of Specialised Translation* 23: 118–36.

Laurenzi, Adrian, Megumu Brownstein, Anne M. Turner, Julie A. Kientz, and Katrin Kirchhoff. 2013. “A Web-Based Collaborative Translation Management System for Public Health Workers.” In *Conference on Human Factors in Computing Systems - Proceedings*, 511–516. DOI 10.1145/2468356.2468446.

Liu, Chao-Hong (ed). 2018. “(LoResMT 2018) Technologies for MT of Low Resource Languages 2018.” In *The 13th Conference of the Association for Machine Translation in the Americas*, 1–62. Boston, Massachusetts, USA: Association for Machine Translation in the Americas.

Liu, Chao-Hong, and Alina Karakanta (eds). 2019. “(LoResMT 2019) The Second Workshop on Technologies for MT of Low Resource Languages.” In *Machine Translation Summit XVII*, 1–77. Machine Translation Summit.

Martindale, Marianna J., and Marine Carpuat. 2018. “Fluency Over Adequacy: A Pilot Study in Measuring User Trust in Imperfect MT.” In *Proceedings of the 13th Conference of the Association for Machine Translation in the Americas (Volume 1: Research Papers)*, 13–25.

Martindale, Marianna, Marine Carpuat, Kevin Duh, and Paul McNamee. 2019. “Identifying Fluently Inadequate Output in Neural and Statistical Machine Translation.” In *Proceedings of Machine Translation Summit XVII Volume 1: Research Track*, 233–243. European Association for Machine Translation.

Matamala, Anna, Andreu Oliver, Aitor Álvarez, and Andoni Azpeitia. 2015. “The Reception of Intralingual and Interlingual Automatic Subtitling: An Exploratory Study within the HBB4ALL Project.” In *Proceedings of the 37th Conference Translating and the Computer*, edited by João Esteves-Ferreira, Juliet Macan, Ruslan Mitkov, and Olaf-Michael Stefanov, 12–17. ASLING.

Matamala, Anna, and Carla Ortiz-Boix. 2016. “Accessibility and Multilingualism: An Exploratory Study on the Machine Translation of Audio Descriptions.” *Trans* 20: 11–24. DOI 10.24310/TRANS.2016.v0i20.2059.

Melby, Alan K, and C Terry Warner. 1995. *The Possibility of Language: A Discussion of the Nature of Language, with Implications for Human and Machine Translation*. Amsterdam/Philadelphia: John Benjamins.

Melero, Maite, Antoni Oliver, and Toni Badia. 2006. “Automatic Multilingual Subtitling in the ETITLE Project.” In *Proceedings of Translating and the Computer 28*, 1–18.

Miyata, Rei, Anthony Hartley, Cécile Paris, Midori Tatsumi, and Kyo Kageura. 2015. “Japanese Controlled Language Rules to Improve Machine Translatability of Municipal Documents.” In *MT Summit XV*, 90–103.

Miyata, Rei, Anthony Hartley, Kyo Kageura, Cécile Paris, Masao Utiyama, and Sumita Eiichiro. 2016. “MuTUAL: A Controlled Authoring Support System Enabling Contextual Machine Translation.” In *COLING 2016*, 35–39.

O’Mathúna, Dónal, Carla Parra Escartín, Helena Moniz, Jay Marlowe, Matthew Hunt, Eric DeLuca, Federico Federici, and Sharon O’Brien. 2019. “Ethics Recommendations for Crisis Translation Settings.” INTERACT: The International Network in Crisis Translation.

Ortiz-Boix, Carla, and Anna Matamala. 2017. “Assessing the Quality of Post-Edited Wildlife Documentaries.” *Perspectives: Studies in Translatology* 25 (4): 571–593. DOI 10.1080/0907676X.2016.1245763.

Parra Escartín, Carla and Helena Moniz. 2020. "Ethical Considerations on the use of Machine Translation and Crowdsourcing in Cascading Crises." In *Translation in*

Cascading Crises, edited by Federico M. Federici and Sharon O'Brien, 132–151. London: Routledge.

Piperidis, Stelios, Iason Demiros, Prokopis Prokopidis, Peter Vanroose, Anja Hoethker, Walter Daelemans, Elsa Sklavounou, Manos Konstantinou, and Yannis Karavidas. 2004. "Multimodal Multilingual Resources in the Subtitling Process." In *Fourth International Conference on Language Resources and Evaluation (LREC 2004)*. ELRA.

Plitt, Mirko, and François Masselot. 2010. "A Productivity Test of Statistical Machine Translation Post-Editing in a Typical Localisation Context." *The Prague Bulletin of Mathematical Linguistics* 93: 7–16. DOI 10.2478/v10108-010-0010-x.

Pouliquen, Bruno, Cecilia Elizalde, Marcin Junczys-Dowmunt, Christophe Mazene, and José Garcia-Verdugo. 2013. "Large-Scale Multiple Language Translation Accelerator at the United Nations." In *Proceedings of the XIV Machine Translation Summit*, 345–351.

Somers, Harold. 1997. "Machine Translation and Minority Languages." In *Translating and the Computer 19: Papers from the Aslib Conference Held on 13 & 14 November 1997*, 1–13. Aslib.

Tiilikka, Päivi. 2013. "Access to Information as a Human Right in the Case Law of the European Court of Human Rights." *Journal of Media Law* 5 (1): 79–103. DOI 10.5235/17577632.5.1.79.

Turner, Anne M., Margo Bergman, Megumu Brownstein, Kate Cole, and Katrin Kirchhoff. 2014. "A Comparison of Human and Machine Translation of Health Promotion Materials for Public Health Practice: Time, Costs, and Quality." *Journal of Public Health Management Practices* 20 (5): 523–29. DOI 10.1016/j.physbeh.2017.03.040.

Turner, Anne M., Yong K. Choi, Kristin Dew, Ming-Tse Tsai, Alyssa L. Bosold, Shuyang Wu, Donahue Smith, and Hendrika Meischke. 2019. "Evaluating the Usefulness of Translation Technologies for Emergency Response Communication: A Scenario-Based Study." *JMIR Public Health and Surveillance* 5 (1): e11171. DOI 10.2196/11171.

UNESCO Executive Board. 2007. “Report by the Director-General on the Execution of the Programme Adopted by the General Conference: Intersectoral Mid-Term Strategy on Languages and Multilingualism.”

United Nations. 1948. “Universal Declaration of Human Rights.”

Vasconcellos, Muriel, and Marjorie León. 1985. “SPANAM and ENGSPAN: Machine Translation at the Pan American Health Organization.” *Computational Linguistics* 11 (2–3): 122–136.

Vasiljevs, Andrejs, Rihards Kalnins, Marcis Pinnis, and Raivis Skadins. 2014. “Machine Translation for E-Government - the Baltic Case.” In *Proceedings of the 11th Conference of the Association for Machine Translation in the Americas, Vol. 2: MT Users Track*, 181–193.

Authors' Addresses

Mary Nurminen
http://orcid.org/0000-0002-2380-6950
Tampere University
Kalevantie 4
Tampere, 33014
Finland
mary.nurminen@tuni.fi

Maarit Koponen
https://orcid.org/0000-0002-6123-5386
University of Turku
School of Languages and Translation Studies
Department of English
Turku, 20014
Finland
maarit.koponen@utu.fi