

“All those in favour will please say yea”: Understanding the Factors Behind Machine Translation Adoption at the Canadian Parliament

Jeniffer Leal-Wyss Delaney Lothian Gabriel Bernier-Colborne
Rebecca Knowles Michel Simard

National Research Council Canada

{Jeniffer.Leal, Delaney.Lothian, Gabriel.Bernier-Colborne,
Rebecca.Knowles, Michel.Simard}@nrc-cnrc.gc.ca

Abstract

Translators at the Canadian Parliament currently have access to a neural machine translation system as an optional tool integrated into their translation environment, whose output they can use for postediting (rather than translating from scratch); this provides a valuable opportunity to study the dynamics of machine translation adoption in professional settings. We report on a user study that investigates how and why translators choose to interact with this tool. Using a mixed-methods approach, we examined both human and technical factors that influence the adoption or non-adoption of the system. Drawing on our findings, we advocate for a user-centred approach to MT integration within professional translation workflows.

1 Introduction

Machine translation (MT) is used in many different contexts for various purposes. One of those is professional translation workflows, where it is used, typically alongside other computer-assisted translation (CAT) tools, for the purpose of postediting (PE). That is the case for the Parliament of Canada, where the transcripts of the House of Commons (HoC) and the Senate must be available in both of Canada’s official languages, French and English. Parliamentary translators have the option to use the Hawkeye MT systems to generate MT output for PE. These are Neural Machine Translation (NMT) systems based on Amazon’s Sockeye toolkit (Hieber et al., 2022), trained

on parliamentary texts, and updated multiple times a year (Knowles et al., 2023; Knowles et al., 2024). Translators can access these systems’ output through their integration into their work environment, the Prism workflow management application (O’Brien, 2002), and can choose whether or not to postedit the MT output.

We present a user study carried out primarily to understand user choices and perceptions of this technology, and to bridge the gap between development and actual usage. This study allows us to understand the system from the perspective of end-users, which can help further improve the system and have a direct impact on the work of parliamentary translators. Furthermore, in many translation workplaces, the use of MT is imposed, and cases where translators within an institution can freely choose whether to use MT or not are relatively rare. Thus, this study also offers a unique opportunity to understand motivations for the adoption or non-adoption of MT from the perspective of users.

The results presented in this document help to answer the following research questions: **(RQ1)** What is parliamentary translators’ experience with using the Hawkeye MT system in their translation work? **(RQ2)** What factors affect the willingness of parliamentary translators to use the Hawkeye MT system? **(RQ3)** How can the Hawkeye MT system’s output and its overall implementation be improved? To answer these questions, the study used a mixed-methods explanatory sequential design, comprising two stages: a quantitative stage and a qualitative stage. The first stage includes a questionnaire collecting data regarding users’ basic experience with and use of the system. The qualitative stage consists of semi-structured interviews with users and is focused on users’ perceptions and requirements of Hawkeye. Partici-

pants were recruited from a pool of 66 parliamentary translators from the Translation Bureau, the federal institution responsible for the translation of government documents into both official languages.

2 Related Work

Our study focuses on MT use for PE at the Parliament of Canada, but it is situated within the broader context of MT use for professional translation, where related research in the field of translation studies can offer valuable insights.

Previous studies about MT use in professional environments have examined adoption and perception of MT, using a variety of methods. For example, Zaretskaya (2015) conducted a survey among translators to identify MT adoption patterns, as well as perceptions of this technology. That study found an overall increase in the adoption of MT compared to previous studies and identified higher usage rates in domains where MT quality was perceived to be better. The author also found that IT literacy and training were factors that had an impact on MT usage.

In another study focusing on the adoption and non-adoption of MT, Cadwell et al. (2018), investigated two distinct groups of users to elucidate the various factors influencing user choices. Through focus group discussions, the authors identified common themes that explain MT adoption choices and the factors influencing those decisions. The study found that some of the most common reasons for the adoption of MT included speed or productivity gains, effective results, good quality, inspiration or new ideas, among others, while the reasons for non-adoption include ineffective results for certain types of texts, poor output quality for certain language pairs, negative influence on the translator's abilities, among others.

Other studies have also used surveys for the purpose of understanding perceptions of translation technologies and MT systems. For instance, Flanagan et al. (2025) conducted a survey to assess adoption of translation technologies among Danish translators, finding that they neither rejected these tools nor embraced them uncritically. Similarly, Moorkens and O'Brien (2013) examined user attitudes towards PE interfaces using a questionnaire, asking participants about their working methods, attitudes towards MT, and their ideal User Interface (UI), among others. This study found the

highest priorities for participants regarding PE UI were customization, a clean and uncluttered view, dynamic changes to MT, and other features.

Other research has used a combination of methods, contributing to a better understanding of PE processes. Heinisch and Lušicky (2019) analyzed expectations towards MT in a case study where they used a mixed-methods approach that included evaluations of raw MT output quality, as well as a questionnaire. Their findings suggest that users' experience and expectations have an impact on the use and evaluation of MT.

A combination of methods has also proven effective in other studies. For instance, Moorkens and O'Brien (2017) examined user perspectives regarding PE interfaces. The authors point out a broader trend in MT research toward prioritizing system development and evaluation over user-centred perspectives. To address this gap, they combined survey data and follow-up interviews with translators to explore user experience in PE, identify interface needs, and suggest potential incorporation of user insights into the UI design process.

3 Motivation and Methodology

Parliamentary translation in Canada is organized into three main areas: Debates and Proceedings of the HoC, Debates and Proceedings of the Senate, and Parliamentary Documents. We focus on the HoC, where debates are usually translated overnight, while committee materials are translated under somewhat less time pressure. The translation process of parliamentary speech involves several stages, of which two are most relevant to this study: translation and revision. Translations are carried out between English and French in both directions. Once completed, translations are sent for revision, where revisers (some of whom are also translators) review, correct, and approve them prior to publication.

To better understand the use of Hawkeye MT systems by translators, we analyzed log files from May 2023 to December 2025 that showed when translators requested MT output for PE for HoC debates and committees texts (Simard et al., 2026). We observed a steady increase in the proportion of texts for which MT was requested, from 16% in May 2023 to approximately 58% in December of 2024, plateauing at about that level throughout 2025. In terms of frequency, most transla-

tors tended to use Hawkeye either seldom or almost always. Additionally, a comparison of revision effort between postedited texts produced with Hawkeye and translations completed without the use of Hawkeye revealed that texts translated with Hawkeye required less revision than those translated without it. However, these usage patterns do not explain *why* translators choose to use (or not use) the available MT systems, which motivated the research questions and approach of the user study presented in this paper.

Since the present study focuses on two different aspects—actual use and perceptions of a specific MT system—we decided to use a mixed-methods explanatory sequential design (Creswell and Creswell, 2017) that comprises two stages: a quantitative stage and a qualitative stage. The initial quantitative stage used a questionnaire to collect data of user’s basic experience with and use of the MT system. Although the questionnaire contained some open text questions, it mainly included multiple-choice questions for quantitative analysis. After the results from the initial stage were analyzed, the second stage, which consisted of semi-structured interviews, was designed with the intent to aid in explaining the results of the initial quantitative stage, by focusing on users’ perceptions and requirements of the Hawkeye system. This approach was chosen because it aligns well with our research objectives and has been demonstrated to be a robust research method in previous studies in the field, as presented above. All participants for both stages were recruited from a pool of 66 parliamentary translators from the Translation Bureau. The described study design was reviewed and approved by the National Research Council of Canada’s Research Ethics Board. The questionnaire and interview protocol included consent forms, and participation was strictly voluntary. Participants were not remunerated for their participation but were permitted to participate in the study during their work hours.

3.1 Stage 1: Questionnaire

The questionnaire¹ consisted of questions concerning respondent experience with the Hawkeye MT tool; specifically regarding the usability of the UI, general experience pertaining to PE and to other CAT tools, as well as some demographic questions

relating to the respondents’ experience in their position. A subset of the questions were mandatory; the remainder were optional. The questionnaire was designed in English and translated into French following established best practices, so it would be available to participants in the language of their choice.

The questionnaire incorporated two standardized instruments: the international positive and negative affect schedule short form (I-PANAS-SF) (Thompson, 2007) and the system usability scale (SUS) (Brooke, 1996). Both of these instruments have corresponding validated French versions. The I-PANAS-SF is used to measure participant affect (i.e., emotion), while the SUS is designed to assess usability. The questionnaire was deployed online by NRC’s Communication Branch, using the Voxco cloud-based survey software platform.

Participants were recruited via email. The recruitment email was written by the research team and was subsequently distributed by the Translation Bureau. The recruitment email included links to both the French and English versions of the questionnaire.

Overall, the participation rate was very positive: of the 66 questionnaire invitations sent, 39 responses were received, representing a total response rate of 59%. Of the 39 responses received, four responses were excluded due to confusion regarding the tool being evaluated, where Hawkeye was confused for an older MT system also available to parliamentary translators in other translation settings. This means the total questionnaire population presented in this work is $n=35$.

Quantitative analysis was conducted using descriptive statistics, and analysis of open answer questions was performed using thematic analysis (Naeem et al., 2023).

3.2 Stage 2: Interviews

This study used semi-structured interviews, where the interviewer asked the participant a set of questions, but was allowed to ask additional questions off-script (e.g., follow-up questions) when deemed necessary (Kallio et al., 2016). This format allowed us to hit key points to ensure we were able to answer our research questions, while also providing flexibility for exploration beyond what we initially anticipated. The interviews were designed to ask participants about their use of and prefer-

¹See Appendix B for the full text of both the questionnaire and interview questions.

ences for other CAT and MT tools, their perception of the quality of the Hawkeye MT system’s output, their motivations for using (or not using) Hawkeye, their experience receiving related training, and their perception of the UI the Hawkeye system is embedded in, as well as their perceptions of PE in general.

Invitation e-mails were sent to participants who volunteered through an optional form at the end of the questionnaire. After confirmation, they received a consent form that was required to be signed before receiving a Microsoft Teams invitation to participate in the online interview. Interviews were offered to participants in either French or English and were led by a primary interviewer with the support of a note-taker. They lasted up to 45 minutes and were all conducted remotely and recorded using the recording tool integrated in Microsoft Teams. After completion of the interviews, they were transcribed using Microsoft’s automatic transcription, followed by human revision. The recordings were destroyed after transcription. Of the 13 participants who initially volunteered to participate in the interviews, we conducted 10 interviews.

Thematic analysis was conducted using an inductive coding process, in which themes were identified and collected from participant open answer style data. Due to the semi-structured nature of the interviews, information relating to each question asked may appear in different places across the interview. Therefore, themes were allowed to be found and grouped across entire interviews, as opposed to question-by-question.

4 Results

In this section, our results are presented in three parts. First, we present the overall response and the results regarding Hawkeye use, which helps us to answer RQ1. Second, participant-reported reasons for using or not using the system, as well as factors considered when deciding whether to use it or not, which helps us answer RQ2. Third, we look at other factors that could potentially have an impact on MT use and examine whether these factors play a role in system adoption, which also helps to answer RQ2. To answer RQ3, we provide our analysis and discussion in Section 5.

4.1 Overall Response and Use of Hawkeye

One of the first questions in our questionnaire asked participants whether they had used the Hawkeye MT system (Question 3, single-response question).² Our results indicate that out of 35 questionnaire respondents, 26 participants (74%) have used Hawkeye, while 9 participants (26%) have not. In the interviews, 9 out of 10 participants (90%) reported having used Hawkeye.

We also asked the 26 questionnaire respondents who reported having used the system about the proportion of their work for which they use it (Question A.2, single-response question). We present the results in Table 1. These findings show that most translators who have tried Hawkeye report using the tool more than half the time, while a minority use it less frequently. In the interviews, 7 out of 10 participants reported using Hawkeye all the time when it is available (i.e., when working for the HoC, as explained in Section 4.2).

% Work	# Users	Percentage
0% (Never use it)	2	8%
1–25%	9	35%
26–50%	0	0%
51–75%	10	38%
76–99%	3	12%
100% (All the time)	2	8%
Total	26	100%

Table 1: Proportion of work for which Hawkeye is used.

We also asked participants about the likelihood that they would continue using Hawkeye in the future (Question A.3, single-response question). Of the 26 participants who reported having used Hawkeye, 25 answered this question, as it was optional. In response, a clear majority (15 out of 25 participants) stated that they were very likely to keep using Hawkeye, 2 out of 25 participants said they were somewhat likely, 2 out of 25 participants said they were not very likely, 3 out of 25 participants said they were not at all likely, and 3 out of 25 participants said they did not know whether they would continue using the Hawkeye MT system.

In this study, we explore user perceptions about MT quality. Quality assessments suggest that overall MT output quality is perceived to be acceptable

²Questions in the questionnaire were formatted as single-response (radio buttons), multi-response (check boxes), matrix (multiple rows of radio buttons or check boxes), and text boxes.

for PE, based on responses to Question A.7 (matrix question). This question asked the 26 participants who reported having used Hawkeye how often *critical*, *major*, *minor*, and *neutral* errors appeared in the MT output. The responses show that *minor errors* were the most frequently reported, with 22 out of 26 participants indicating they appeared “sometimes” or “often,” followed by *neutral errors* with 21 out of 26 participants reporting the same. In contrast, *major errors* were less common, with 13 participants noting they appeared “sometimes” and none reporting “often.” *Critical errors* were the rarest, with only 2 participants indicating they appeared “sometimes” and none reporting “often.”

In this context, we also asked participants whether they had noticed any changes in quality over time, since Hawkeye is updated 3 times per year (Question A.4, single-response question). A plurality of respondents (11 out of 26 participants) reported perceived improvements, while a smaller proportion (3 out of 26 participants) felt that quality had declined. A similar share (4 out of 26 participants) indicated that they felt the quality had remained the same, and 8 out of 26 participants said they were unsure.

Our results suggest that participants find the UI simple and easy to use, as reported by responses to the System Usability Scale (Question A.10, matrix question), answered by the 26 participants who reported having used Hawkeye. This scale is a standardized usability instrument that provides a frame of reference and benchmark for understanding how an interface compares against the average. For context, it has been found that 68 is the median SUS score. The SUS for the implementation of the Hawkeye system in the Prism UI was 76.35 (SD=18.94) with a 95% CI = [72.88, 79.81]. Based on the “grade” and “adjective” ratings (“worst imaginable”, “poor”, “OK”, “good”, “excellent”, “best imaginable”) established for the SUS (Sauro and Lewis, 2016), this score puts the Hawkeye UI in the B-grade and “good” ranges, respectively.

4.2 Participant-Stated Reasons for the Adoption and Use of Hawkeye

We now present the different reasons that might explain why participants choose to use or not to use the Hawkeye MT system, as well as the factors that influence their choice. First, we present participant-reported factors that influence adoption

and use.

We asked the 26 questionnaire participants who reported having used Hawkeye about their reasons for doing so (Question A.B.1, multi-response question); 24 participants answered this optional question. Our results show that people use it mainly as a jumping-off point/source of inspiration (22/24), to work faster (11/24), to reduce typing (11/24), to automate the most repetitive parts (10/24), for other reasons (8/24), and to reduce mental effort (6/24) (see Figure 1).

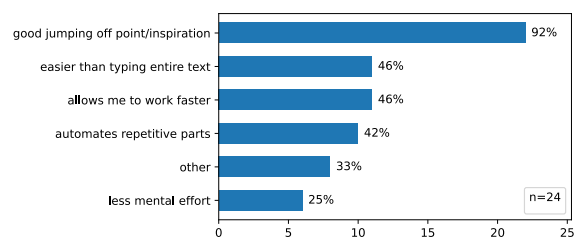


Figure 1: Question A.B.1 - I use the Hawkeye translation tool because (Select all that apply)

Among the participants who reported using Hawkeye for other reasons, the most common reason provided was ergonomics, with 3 out of 8 participants explaining that MT helped them reduce typing and, therefore, the physical stress of translating. This finding is noteworthy, as “reduce typing” was already an option provided in the questionnaire, but the fact that the participants elaborated on this offers an unexpected insight into the ergonomics of professional translation: the physical impact of typing and how the use of MT can help reduce this strain.

Another relevant factor to determine was why some translators chose to use Hawkeye for a portion of their work. To explore this, we asked participants who reported using Hawkeye consistently (n=24, a subset of the 26 who have tried it at least once) about the factors they consider when deciding whether or not to use the system (Question A.B.2, open text question). This optional question was answered by 16 participants. The most frequent response (9 out of 16 participants) was that the determining factor was the client, i.e., whether they are working for the HoC or the Senate. To understand this trend, it is important to clarify that, when the questionnaire and interviews were conducted, the Senate’s Hawkeye systems had not yet been made available to parliamentary linguists. Since translators work for both the HoC and the Senate, their option to use

it is constrained by the source of the text they are translating. If the client is the Senate, Hawkeye is not available to them. Other factors that influenced whether linguists used Hawkeye included the type of content to be translated (3 out of 16 participants), as well as system performance (3 out of 16 participants), which includes technical problems such as freezing, slow loading times, or crashing, and appears to negatively impact both the ability and motivation to use the tool.³ In addition, participants mentioned other factors, such as tighter deadlines, assessing output quality to determine usability, formatting, effort, whether they are working on Prism, and the availability of alternative resources, although they didn't always explain whether these factors meant using MT or not.

There were 2 participants who had tried Hawkeye but did not use it on a regular basis; we also asked them about their reasons for not doing so (Question A.A.1, open text question). One participant noted that, as a reviser,⁴ they did not need to use MT, while another explained that they did not consider it beneficial in terms of quality or speed.

We also asked participants in our questionnaire who reported that they had never used Hawkeye the reasons why they do not use it (Question B.1, open text question, see 2). Of the 9 participants who reported not using Hawkeye, 8 participants answered this optional question, and some of the answers were categorized in more than one theme during thematic analysis: 3 participants noted that they prefer to use DeepL,⁵ an MT system that is available to all translators, but that is not integrated into their work environment, 2 participants noted they lacked the resources to use it (one of them in terms of training and the other one noting they didn't have the tool installed in their computer, which also suggests that they might have knowledge gaps on how to use it), 2 participants stated that they didn't like to use MT, 2 participants reported they did not believe it would help them save time or reduce effort, and 2 participants expressed quality concerns. One participant also noted that the system often had issues, and another explained that MT was not needed for revision work.

The results of this section are particularly inter-

³At the time of writing, the system latency and crashing issues have been improved by increasing computing capabilities within this workflow. Additionally, the Senate systems have been made available to translators.

⁴A small number of participants in the study perform only revision work, without any translation or postediting tasks.

⁵<https://www.deepl.com/en>

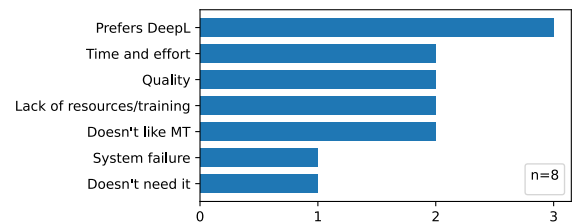


Figure 2: Question B.1 (for participants who do not use Hawkeye) - Why do you not use the Hawkeye translation tool?

esting because they highlight that not using Hawkeye does not necessarily mean not using MT. As noted in the questionnaire responses, participants have institutional access to DeepL Pro, and some prefer to use that tool instead of Hawkeye. We explored this further during the interviews, asking participants if they used another MT system instead of Hawkeye and what the benefits of using the other tool were (Interview Question 3: *Have you ever used other machine translation tools different from Hawkeye for your work at the parliamentary service?* and subsequent follow up question: *When you are using the other MT tool instead of Hawkeye, what is something that you can do that you could not do with Hawkeye?*). Out of 10 interviewees, 3 explained that they use DeepL for Senate content because Hawkeye is not available for that content; another 3 stated that they use DeepL because it offers word alternatives for MT output; 2 said they use DeepL when they are not satisfied with Hawkeye's output quality; another reported using it when Hawkeye had system issues.

Additionally, comparisons of output quality between Hawkeye and DeepL noted that Hawkeye produced better output for parliamentary content (6 out of 10 interview participants), while others indicated that DeepL performed better for general language (4 out of 10 interview participants), a difference that is likely to be linked to the data used for system training. Other remarks comparing DeepL to Hawkeye also explained that DeepL's output uses European French, while Hawkeye is more appropriate for a Canadian audience.

4.3 Other Factors Potentially Affecting Adoption and Use of Hawkeye

After examining participants' reported reasons for using or not using Hawkeye, we turned to other factors that may influence system usage. Prior work (Moorkens and O'Brien, 2015) has found demographic attributes such as years of translation

experience to be correlated with use of MT. However, the user population in our study is quite experienced (no participants with fewer than 10 years of experience in the language industry; Question 14, single-response question) and most have long tenure in the parliamentary service (20 of 35 respondents reported more than 10 years of experience with the Prism environment; Question 2, single-response question). Within our small and notably experienced cohort, we did not find clear demographic correlations with Hawkeye use, with the exception of a small number of cases where participants serve primarily as revisers rather than translators (though this has more to do with the particular responsibilities of their role).

Another factor we examined was the assumption that adequate training in the use of MT as a translation aid encourages its adoption. To explore this, we asked translators whether they had received training in the use of translation tools (Question 4, matrix question). We found that 28 of the 35 respondents (80%) had received training in “MT tools and PE”. Of these 28, 11 use Hawkeye regularly, compared to 4 of 7 respondents without MT training who use Hawkeye regularly. Expanding the scope to include respondents who either use Hawkeye regularly or reported using another MT system in Question 7, we observe that 19 of the 28 trained in MT use some form of MT, compared to 5 of 7 not trained in MT who use some MT. These findings suggest that there is no clear correlation between training in the use of MT and PE, and the actual use of Hawkeye. Furthermore, during the interviews, we asked participants whether training affected their willingness to use the Hawkeye MT system (Interview Question 9.A: *Do you think that receiving appropriate training affects your willingness to use the Hawkeye machine translation system?*), to which 8 of 10 participants said “no”.

Since some participants noted in the UI evaluation section of the questionnaire that they lacked training on how to use the Hawkeye system, we decided to expand on this question in the interviews, asking participants whether they had received training related to Hawkeye, MT, or PE (Interview Question 9: *Could you please tell me about any training related to machine translation, postediting and the Hawkeye machine translation system that you have received at your current workplace?*). A large majority (9 out of 10 interview participants) answered they had not re-

ceived any training specific to Hawkeye.

During the interviews, we also asked participants what they would like to learn in training to help them use the MT system effectively (*Question 9b: What would you like to learn during training that you think would help you to use the machine translation system?*). 4 out of 10 participants noted they would like to receive training on basic functionalities, UI buttons, or system customization. Additionally, two participants expressed interest in learning about key considerations for PE and understanding the system’s limitations, and two other participants mentioned that revision training was useful for PE as well. Some participants also expressed an interest in learning about artificial intelligence. These findings show that there is no clear correlation between MT and PE training and the adoption of Hawkeye, but that there may be interest in training for other reasons.

Another factor we explored was whether familiarity with other CAT tools affected translator’s willingness to use Hawkeye. We asked participants which other translation tools they use (*Question 7, open text question*). Of the 35 respondents, 22 (63%) reported using MT systems other than Hawkeye. In most cases, this “other MT” was DeepL, but some also reported using other custom-built NMT solutions provided by the Translation Bureau. In some open text questions (*Question A.14; Question B.1*), some respondents suggested that they preferred DeepL over Hawkeye. However, in practice, 13 of the 22 respondents who reported using other MT systems were also regular users of Hawkeye.

Putting these numbers together, we get the following picture: among our population of 35 translators, 24 use MT to translate parliamentary texts (69%). Of these, 15 report using Hawkeye regularly (63% of MT users; 43% overall); and 9 use other MT systems, most likely DeepL (37% of MT users, 26% overall), although we do not know if they use those systems regularly. Another noteworthy aspect is that, of these 9 participants who reported using other MT systems and who are not regular users of Hawkeye, 5 had never actually tried Hawkeye (as reported in Question 3). This leaves only 4 translators who have tried Hawkeye and still prefer to use DeepL.

An additional factor we explored was whether experience with PE affected adoption of Hawkeye. We asked participants whether they had ever

worked in PE (Question 8, single-response question); of our population of 35 respondents, 34 answered this question, where 21 participants replied “no” (62%), and 13 participants answered “yes” (38%). This finding was somewhat unexpected, since, as mentioned earlier, a majority of participants use Hawkeye or other MT tools in their work. To better understand this discrepancy, we examined the topic in greater depth during the interviews, where only 2 of 10 interview participants reported that they considered themselves posteditors. In this sense, some translators appeared to be unsure of the terms *posteditor* and *postediting*, and others noted that they were unfamiliar with the specific tasks involved in PE, that PE was less interesting than translation, or that MT is simply a tool that supports their work without changing the nature of their role, while still others explained that they did not fully consider themselves posteditors because they use MT for only part of their work. The discrepancy and the semantic nuance observed in these results limits our ability to further investigate the potential correlation between PE experience and the use of Hawkeye.

One final factor is that of personal preference. We asked translators which type of translation activity they preferred (Question 10, single-response question). Our results show that 16 out of 35 participants prefer translating from scratch over PE, 13 out of 35 participants prefer PE over translation, 4 out of 35 participants enjoy both tasks equally, and 2 out of 35 participants were not sure.

This finding becomes particularly significant when analyzed in relation to Question 3 (*Have you ever used the Hawkeye machine translation tool integrated in Prism for your work in the Parliamentary Translation Service*) and Question A.2. (*For what proportion of your work do you estimate that you currently use the Hawkeye translation tool?*). Among the 16 participants who prefer translating from scratch, 5 have never tried Hawkeye, 2 do not use it regularly, and 6 use it only rarely. Only 3 participants who prefer translating from scratch reported using Hawkeye regularly. This suggests that personal preference between translation and PE is an important adoption factor, one that extends beyond the system itself.

5 Discussion

The adoption and non-adoption of the Hawkeye NMT system by parliamentary translators offers

a valuable opportunity to understand MT adoption patterns from the perspective of professional translators who use it for PE. In many translation workplaces, the use of MT is imposed (Vieira and Alonso, 2020), and the cases where translators within an institution can freely choose whether or not to use MT are relatively rare. Thus, this study offers a unique opportunity to understand motivations for the adoption or non-adoption of MT from the perspective of users.

When analyzing the results of our study, we find trends of increasing MT adoption. Firstly, results from log file analysis show an increase in Hawkeye use. Secondly, our survey found that a majority of respondents use Hawkeye and/or other MT tools and that they are very likely to continue to use them. In contrast to other work (O’Brien and Moorkens, 2014), our survey did not identify a generalized trend of resistance to the adoption of MT. Below, we offer an analysis of the different reasons why translators may or may not adopt MT for PE at the Parliament of Canada.

According to the Technology Acceptance Model (Davis, 1989), users adopt technologies primarily for two reasons. The first is perceived usefulness, which in the case of MT may involve faster turnaround times, thus improving productivity, as well as improved translation quality, for example, through more accurate specialized terminology. The second is perceived ease of use, which may include having a simple UI or seamless integration of the MT system into the existing work environment, thereby reducing the effort involved in switching between windows to access MT.

The results of our study show that translators who adopt Hawkeye generally do so because they find it useful. Some of the most obvious advantages included: to work faster, to reduce mental effort, to reduce typing, and to automate the most repetitive parts of the work. However, the primary reason reported for using Hawkeye was that it served as a jumping-off point or source of inspiration, which was somewhat unexpected, as various studies on postediting emphasize productivity as the main advantage of using MT (Guerberof, 2009; Stasimioti and Sosoni, 2020; Macken et al., 2020), an option that was offered to participants among the possible alternatives. One could argue, of course, that having a starting point may also contribute to working more quickly, since evaluating a proposed solution often takes less time than

generating one from scratch. However, when considering the underlying benefits of using MT as a source of inspiration, these benefits appear to be more related to reducing cognitive effort and finding better solutions, rather than increasing working speed. This was further illustrated by one of our interview participants, who stated:

“I don’t see it as a replacement, but rather as a helpful tool that provides clues and solutions. Sometimes, it suggests beautiful words I wouldn’t have thought of myself” (English translation, Interview Participant 8).⁶

Another element that is important to highlight here is that some participants also mentioned ergonomic reasons as one of the main motivations for using Hawkeye. This was not a factor considered in the design of our study, but participants’ insights provided a deeper understanding of health risks associated with professional translation and how the use of translation tools can help to reduce the physical constraints that are typically associated with extensive typing, as shown in this quote:

“Because of physical constraints, it is not possible for me to type the large number of words that I have to translate everyday. It is for that reason that I now use Hawkeye” (English translation, Questionnaire Participant 9).

Other questionnaire remarks also mentioned factors related to ease of use, such as the fact that the MT system is integrated in their work environment, that it does the formatting for them, etc.

When we examine the reasons why some translators use the tool for only part of their work, two main external factors emerge. The first one involves the client: Hawkeye had not yet been deployed for the Senate, which prevented users from accessing the system. The second one is system issues, which prevented users from working with the tool. If we exclude external factors that are out of the users’ control, we find other factors influencing the decision on when to use Hawkeye. For instance, some participants noted that the content of the source text played a role in this choice, or that they preferred to use a different MT system.

As some participants reported a preference for using DeepL over Hawkeye, we explored how these tools compared further during the interviews, where some participants noted that they appreciated the feature allowing them to select alternative translations when dissatisfied with the initial

output (whereas Hawkeye produces only a single static translation). Other remarks highlighted perceived differences in output quality between Hawkeye and DeepL, noting that Hawkeye’s output was considered superior for specialized parliamentary language, whereas DeepL performed better with general or informal language, but occasionally produced vocabulary reflecting European rather than Canadian French usage. This suggests that for certain types of content, translators might perceive that it is easier to use another MT system or to translate from scratch, as low quality output can significantly increase the time it takes to post-edit MT. This observation also connects to other comments indicating that output quality was a key factor in the decision to use or not use Hawkeye, which is consistent with the reasons discussed earlier.

Additionally, some participants who do not adopt Hawkeye reported that they prefer to use a different tool (in most cases, DeepL), that they don’t believe it would make them work faster or better, that the output quality was not considered to be good enough, or that they hadn’t received training on how to use the system.

In our study, we identified that a majority of translators who do not adopt the system are participants who report disliking MT, an inclination linked to conceptions about the technology more generally rather than specifically related to Hawkeye’s functionality, interface, or output quality. Additional reasons for non-adoption included external factors such as Hawkeye’s unavailability for Senate work, and occasional technical issues like system freezes or crashes that hindered its use. Additional external factors included the non-applicability of MT to revision tasks.

However, we also identified other factors that show potential for improvement. One of them is training. Even if the number of participants who reported having gaps in their knowledge of how to use Hawkeye was a small proportion, it is believed that adequate preparation of those participants could further increase Hawkeye adoption. Furthermore, several interviewees mentioned that they had received training about MT errors and they would like to receive such training in the future, which could suggest that familiarity with MT technology might play a role in MT adoption. Based on the assessment of training needs, which concluded that most participants found the system

⁶See Appendix A for original French versions of quotes.

simple and easy to use, it is recommended that users who identify knowledge gaps on the use of Hawkeye be provided with a concise user guide, cheat sheet, or video tutorial rather than a full training program. Additional training in MT errors, capabilities, and PE processes could also be valuable.

In some cases, participants reported perceiving little or no improvement in quality or productivity when postediting MT output. Several factors may account for this perception. For example, prior research has shown that perceived productivity gains do not always correspond to actual productivity gains in PE (Gaspari et al., 2014). Moreover, studies suggest that PE efficiency often increases after an initial learning period (Groves and Schmidtke, 2009). Strengthening collaboration between research and professional practice could help clarify tangible benefits of PE and potentially support its adoption. As our study design does not allow us to measure changes in productivity, we leave such analysis to future work.

A third factor that may further support MT adoption is continued improvement in output quality. Several participants expressed distrust in the adequacy of MT output for PE, suggesting that higher quality could encourage greater use. Hawkeye is regularly retrained with updated parliamentary data, and quality enhancements are expected over time—an observation consistent with participants' reports of gradual improvement. When discussing MT errors, participants noted issues typical of such systems, including difficulties with informal or idiomatic language, figurative expressions, long sentences, vague discourse, co-reference, proper names, hallucinations, etc. While some of these limitations are inherent to MT, others, such as co-reference resolution, could be investigated and may be possible to improve on (e.g., through the use of larger context windows or with external information sources).

Lastly, interview feedback suggested that enhancements to the UI and underlying systems could also encourage adoption. Some participants highlighted the usefulness of features such as an alternative translation feature, as offered by DeepL, indicating that UI elements designed to improve efficiency may further increase adoption.

In all cases identified as having potential for improvement, the factors involved point to the need for a user-centred approach where developers, managers, and users share system knowledge

to strengthen user preparedness and improve system usability and performance.

6 Conclusion

To answer our RQ1, this study concludes that the MT system deployed for translation at the Parliament of Canada, Hawkeye, offers significant value to many parliamentary translators as a translation aid by providing a jumping-off point or inspiration, generating useful translation solutions, and offering ergonomic benefits through reduced typing effort, among others.

Addressing our RQ2, this study found no evidence that professional experience in the translation industry, parliamentary service, or PE, nor training in PE or MT tools, influenced MT adoption. In contrast, the adoption of Hawkeye appeared to be influenced primarily by personal preferences, such as translation-from-scratch or the use of alternative tools, like DeepL. Other factors identified in this study are external and out of the users' control, such as system availability, and system performance, while others are related to specific work tasks, such as revision. Additional factors discouraging MT adoption are more nuanced and require closer attention, such as training, a deeper understanding of MT output errors, and the integration of additional usability features to support system improvement. These are identified as areas that could potentially enhance adoption. Together, improved user preparedness and stronger output quality and system usability are expected to increase trust in the system, improve the perceived usefulness of Hawkeye, and potentially drive broader adoption over time.

Lastly, in response to RQ3, participants highlighted several UI features—such as pre-population and deletion buttons—and MT issues, including co-reference and terminology errors, that warrant further investigation and, if implemented, could lead to improvements in Hawkeye's output and functionality.

Acknowledgments

We wish to thank the parliamentary translation team and all our partners at the Canadian government's Translation Bureau, without whom this work would not have been possible. We thank Kerry Butt for his assistance in implementing and deploying the questionnaire.

References

- Brooke, John. 1996. SUS—a quick and dirty usability scale. In *Usability evaluation in industry*, pages 189–194. Taylor & Francis.
- Cadwell, Patrick, Sharon O’Brien, and Carlos Teixeira. 2018. Resistance and accommodation: factors for the (non-) adoption of machine translation among professional translators. *Perspectives*, 26(3):301–321.
- Creswell, John W and J David Creswell. 2017. *Research design: Qualitative, quantitative, and mixed methods approaches*. Sage publications.
- Davis, Fred. 1989. Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3):319–340.
- Flanagan, Marian, Helle Jensen, Kristine Bundgaard, and Tina Christensen. 2025. Technology adoption among Danish translators: Practices, perceptions and prospects. *Tradumàtica tecnologies de la traducció*, pages 111–136, 12.
- Gaspari, Federico, Antonio Toral, Sudip Kumar Naskar, Declan Groves, and Andy Way. 2014. Perception vs. reality: measuring machine translation post-editing productivity. In O’Brien, Sharon, Michel Simard, and Lucia Specia, editors, *Proceedings of the 11th Conference of the Association for Machine Translation in the Americas*, pages 60–72, Vancouver, Canada, October 22–26. Association for Machine Translation in the Americas.
- Groves, Declan and Dag Schmidtke. 2009. Identification and analysis of post-editing patterns for MT. In *Proceedings of Machine Translation Summit XII: Commercial MT User Program*, Ottawa, Canada, August 26–30.
- Guerberof, Ana. 2009. Productivity and quality in MT post-editing. In *Beyond Translation Memories: New Tools for Translators Workshop*, Ottawa, Canada, August 26–30.
- Heinisch, Barbara and Vesna Lušicky. 2019. User expectations towards machine translation: A case study. In Forcada, Mikel, Andy Way, John Tinsley, Dimitar Shterionov, Celia Rico, and Federico Gaspari, editors, *Proceedings of Machine Translation Summit XVII: Translator, Project and User Tracks*, pages 42–48, Dublin, Ireland, August. European Association for Machine Translation.
- Hieber, Felix, Michael Denkowski, Tobias Domhan, Barbara Darques Barros, Celina Dong Ye, Xing Niu, Cuong Hoang, Ke Tran, Benjamin Hsu, Maria Nadejde, Surafel Lakew, Prashant Mathur, Anna Currey, and Marcello Federico. 2022. Sockeye 3: Fast neural machine translation with pytorch.
- Kallio, Hanna, Anna-Maija Pietilä, Martin Johnson, and Mari Kangasniemi. 2016. Systematic methodological review: Developing a framework for a qualitative semi-structured interview guide. *Journal of advanced nursing*, 72(12):2954–2965.
- Knowles, Rebecca, Samuel Larkin, Marc Tessier, and Michel Simard. 2023. Terminology in neural machine translation: A case study of the Canadian Hansard. In Nurminen, Mary, Judith Brenner, Maarit Koponen, Sirkku Latomaa, Mikhail Mikhailov, Frederike Schierl, Tharindu Ranasinghe, Eva Vanmassenhove, Sergi Alvarez Vidal, Nora Aranberri, Mara Nunziatini, Carla Parra Escartín, Mikel Forcada, Maja Popovic, Carolina Scarton, and Helena Moniz, editors, *Proceedings of the 24th Annual Conference of the European Association for Machine Translation*, pages 481–488, Tampere, Finland, June. European Association for Machine Translation.
- Knowles, Rebecca, Samuel Larkin, Michel Simard, Marc A Tessier, Gabriel Bernier-Colborne, Cyril Goutte, and Chi-kiu Lo. 2024. Some tradeoffs in continual learning for parliamentary neural machine translation systems. In Knowles, Rebecca, Akiko Eriguchi, and Shivali Goel, editors, *Proceedings of the 16th Conference of the Association for Machine Translation in the Americas (Volume 1: Research Track)*, pages 102–118, Chicago, USA, September. Association for Machine Translation in the Americas.
- Macken, Lieve, Daniel Prou, and Arda Tezcan. 2020. Quantifying the effect of machine translation in a high-quality human translation production process. *Informatics*, 7(2).
- Moorkens, Joss and Sharon O’Brien. 2013. User attitudes to the post-editing interface. In O’Brien, Sharon, Michel Simard, and Lucia Specia, editors, *Proceedings of the 2nd Workshop on Post-editing Technology and Practice*, Nice, France, September 2.
- Moorkens, Joss and Sharon O’Brien. 2015. Post-editing evaluations: Trade-offs between novice and professional participants. In El-Kahlout, İlknur Durgar, Mehmed Özkan, Felipe Sánchez-Martínez, Gema Ramírez-Sánchez, Fred Hollowood, and Andy Way, editors, *Proceedings of the 18th Annual Conference of the European Association for Machine Translation*, pages 75–81, Antalya, Turkey, May.
- Moorkens, Joss and Sharon O’Brien. 2017. Assessing user interface needs of post-editors of machine translation. In Kenny, Dorothy, editor, *Human Issues in Translation Technology*, pages 109–130. Routledge.
- Naeem, Muhammad, Wilson Ozuem, Kerry Howell, and Silvia Ranfagni. 2023. A step-by-step process of thematic analysis to develop a conceptual model in qualitative research. *International Journal of Qualitative Methods*, 22:1–18.

- O'Brien, Sharon and Joss Moorkens. 2014. Towards intelligent post-editing interfaces. In *Proceedings of the 20th FIT world congress*, pages 131–137, August.
- O'Brien, Audrey. 2002. Prism: The house of commons integrated technology project. *Canadian Parliamentary Review*, 25(2):20–22.
- Sauro, Jeff and James R Lewis. 2016. *Quantifying the user experience: Practical statistics for user research*. Morgan Kaufmann.
- Simard, Michel, Jeniffer Leal-Wyss, Gabriel Bernier-Colborne, and Rebecca Knowles. 2026. A longitudinal study of the adoption of specialized MT systems in Canadian parliamentary translation. In *Proceedings of the 26th Annual Conference of the European Association for Machine Translation*.
- Stasimioti, Maria and Vilemini Sisoni. 2020. Translation vs post-editing of NMT output: Insights from the English-Greek language pair. In Ortega, John E., Marcello Federico, Constantin Orasan, and Maja Popovic, editors, *Proceedings of 1st Workshop on Post-Editing in Modern-Day Translation*, pages 109–124, Virtual, October. Association for Machine Translation in the Americas.
- Thompson, Edmund R. 2007. Development and validation of an internationally reliable short-form of the positive and negative affect schedule (PANAS). *Journal of cross-cultural psychology*, 38(2):227–242.
- Vieira, Lucas Nunes and Elisa Alonso. 2020. Translating perceptions and managing expectations: An analysis of management and production perspectives on machine translation. *Perspectives*, 28(2):163–184.
- Zaretskaya, Anna. 2015. The use of machine translation among professional translators. In *Proceedings of the EXPERT Scientific and Technological Workshop*, pages 1–12.

A Original Quotes

In this section, we provide for reference the French and English versions of the quotes presented earlier in the paper.

Quote 1:

“Je ne vois pas ça comme un remplaçant mais comme un outil d’aide qui donne des pistes, des solutions. Puis, des fois je trouve quand même qu’ils utilisent des vraiment beaux mots auxquels, que je n’aurais pas pensé à utiliser” (French original quote, Interview Participant 8)

“I don’t see it as a replacement, but rather as a helpful tool that provides clues and solutions. Sometimes, it suggests beautiful words I wouldn’t have thought of myself” (English translation, Interview Participant 8)

Quote 2:

“En raison de problèmes physiques, il n’est pas possible pour moi de taper quotidiennement le nombre élevé de mots que j’ai à traduire. C’est pour ces raisons que j’utilise maintenant Hawkeye” (French original quote, Questionnaire Participant 9)

“Because of physical constraints, it is not possible for me to type the large number of words that I have to translate everyday. It is for that reason that I now use Hawkeye” (English translation, Questionnaire Participant 9)

B Questionnaire and Interview Text

Here we reproduce the full text of our questionnaire and interview questions.

Introduction

Thank you for participating in this study.

In this questionnaire, you are invited to share your experience of using technologies, and specifically **machine translation (MT) tools**¹ in your work. There are no right or wrong answers to the questions; we are only interested in your experience, in gauging the usefulness of machine translation tools for you, and in improving their effectiveness to meet your needs.

The questionnaire includes questions about your experience, your role, and your relevant background, so that we can better understand your needs. It also asks about your experience with various **translation tools**,² as well as your perceptions of these tools and how they are integrated in your workflow. The questionnaire should take about 30 minutes to complete. If you need to take a break, you can close the survey and return to complete it later, using the link you received in your invitation email.

¹ **Machine translation (MT) tools** are systems such as Google Translate or the Hawkeye system integrated in Prism, which automatically render a text in one natural language into another.

² **Translation tools**: software tools that are used to assist human translators in their work (also known as computer-assisted translation tools, or CAT tools). Some examples of these might include: machine translation tools, concordancers, translation memory systems, terminology management tools, Generative AI or Large Language Model-based tools.

Informed Consent

Having read the Informed Consent document [link], I consent to participate in this study (Mandatory question). [radio buttons]

Yes

No

[skip logic – end]

1 Technology use

1. Do you work with the Prism software system, specifically the translation module (which we will refer to as "Prism" for the rest of this questionnaire)? (Mandatory question). [radio buttons]

Yes

No

[skip logic - exit]

2. How long have you been working with Prism? (Optional question). [radio buttons]

Less than 2 years

2 to 5 years

6 to 10 years

More than 10 years

I don't know

3. Have you ever used the Hawkeye machine translation tool integrated in Prism for your work in the Parliamentary Translation Service (by requesting and/or viewing the output of the MT tool for one or more segments, whether or not you ultimately use any of it)? (Mandatory question). [radio buttons]

Yes, I have

No, I haven't

[skip logic – sections A (yes) and B (no)]

A. Section for users of machine translation in Prism

A.1. When did you most recently use the Hawkeye translation tool? (Optional question).

[radio buttons]

Today

In the last week

In the last month

In the last 6 months

Longer than 6 months ago

I don't know

A.2. For what proportion of your work do you estimate that you currently use the Hawkeye translation tool? (Optional question). [radio buttons]

0% (I don't use the Hawkeye translation tool)

1% to 25%

26% to 50%

51% to 75%

76% to 99%

100%

I don't know

[Skip logic – Sections AA (0%) and AB (all other responses)]

A.A.1. Why do you not use the Hawkeye translation tool? (Optional question). [Text box]

A.B.1. I use the Hawkeye translation tool because (Select all that apply): (Optional question). [checkboxes]

It allows me to work faster

It takes less mental effort to work with it than without it

It gives me a good jumping off point/inspiration

It automates the most repetitive parts

It is easier to make small corrections than typing the entire text

Other (Please specify)

A.B.2. Please list up to three factors you consider when deciding whether to use the Hawkeye translation tool for a given task? (Optional question). [text box]

A.B.3. When the output of the Hawkeye translation tool is partially but not entirely acceptable, how do you work with it? (Select all that apply) (Optional question).

[checkboxes]

Replace only the parts of the segment that need to be changed, keeping as much as possible

Replace longer parts of the segment, even if it means retyping some correct parts

Delete the full segment and retype the translation, using the machine translation as inspiration

Other (please specify)

Back to common A section

A.3. In the future, how likely are you to use the Hawkeye translation tool on a consistent basis? (Optional question). [radio buttons]

Very likely

Somewhat likely

Not very likely

Not at all likely

I don't know

Machine Translation Use

This section will ask about two elements of your experience with machine translation: your assessment of output quality, and your evaluation of the interface and your **user experience**.³ Each of these has its own set of questions below.

³ **User experience** refers to the full range of user perceptions of and reactions to products such as software tools. The user experience begins with the first introduction or exposure to the tool, and encompasses all of the different ways that you interact with or think about the tool from that point on.

Machine Translation Quality Assessment

A.4. Over the time you have used the Hawkeye translation tool, the quality of its output has: (Optional question). [radio buttons]

Obviously decreased in quality

Slightly decreased in quality

Remained unchanged

Slightly improved in quality

Obviously improved in quality

I am not sure

A.5. When using the Hawkeye translation tool, what types of errors do you usually observe, and how frequently do you observe them? (Mandatory question). [matrix question]

Error type	Never	Rarely	Sometimes	Often	I don't know
Terminology errors					
Accuracy (such as omissions, additions, mistranslations)					
Linguistic conventions (such as grammar, spelling, punctuation, spacing)					
Style (such as content that deviates from the organizational style guides or exhibits inappropriate register)					
Locale conventions (such as units of measurement, currency)					
Audience appropriateness (translation invalid or inappropriate for the readers)					
Design and markup (such as tags)					
Other					

A.6. Please indicate up to three error types from the following list that you would most like to see eliminated from the Hawkeye translation tool's output. (Optional question). [Checkboxes]

Terminology errors

Accuracy (such as omissions, additions, mistranslations)

Linguistic conventions (such as grammar, spelling, punctuation, spacing)

Style (such as content that deviates from the organizational style guidelines or exhibits inappropriate register)

Locale conventions (such as units of measurement, currency)

Audience appropriateness (translation invalid or inappropriate for the readers)

Design and markup (such as tags)

Other (please specify):

Below are descriptions of error severity.

- **Neutral** errors are not wrong *per se*, but do not match the user's expectations or preferences.
- **Minor** errors are errors that violate language conventions or instructions, or do not entirely or precisely convey the desired meaning, but do not interfere with comprehension of the message.
- **Major errors** violate language conventions or fail to convey the desired meaning to the extent that they interfere with comprehension of the message.
- **Critical errors** are those that are severe enough to impact things such as the safety of the reader, the ultimate goal of the text, or the reputation of the organization.

A.7. Based on the above defined error severities, please indicate the frequency of error severity occurrence in the Hawkeye translation output. (Mandatory question). [matrix question]

Error severity	Never	Rarely	Sometimes	Often	I don't know
Neutral					
Minor					
Major					
Critical					

Please note that the following question uses an existing standardized list of emotions and was not created specifically for this questionnaire. Using this standardized list allows us to generalize the results.

A.8. Think about when you encounter an error in the Hawkeye translation output. To what extent do you generally feel: (Mandatory question). [matrix question]

	(Never) 1	2	3	4	(Always) 5
Upset					
Hostile					
Alert					
Ashamed					
Inspired					
Nervous					
Determined					
Attentive					
Afraid					
Active					

A.9. Is there anything else you would like to say about the Hawkeye translation tool's output? (Optional question) [Text box]

User interface and experience

A.10. The following questions are intended to assess your user experience with the user interface of the Hawkeye translation tool within Prism. Please focus not on the output that is produced, but rather on the way(s) you interact with the tool (Mandatory question).
[matrix question]

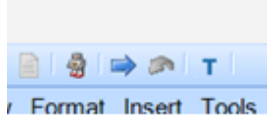
The system usability scale standard version

Strongly agree Strongly disagree

1 2 3 4 5

1	I think that I would like to use the Hawkeye user interface frequently					
2	I found the Hawkeye user interface unnecessarily complex					
3	I thought the Hawkeye user interface was easy to use					
4	I think that I would need the support of a technical person to be able to use the Hawkeye user interface					
5	I found the various functions in the Hawkeye user interface were well integrated					
6	I thought there was too much inconsistency in the Hawkeye user interface					
7	I would imagine that most people would learn to use the Hawkeye user interface very quickly					
8	I found the Hawkeye user interface very awkward to use					
9	I felt very confident using the Hawkeye user interface					
10	I needed to learn a lot of things before I could get going with the Hawkeye user interface					

A.11. You can access the output from the Hawkeye translation tool in Prism in different ways. Which one(s) do you commonly use? (Optional question) [checkboxes]



[Image description: A screenshot of the “Robot” and the “T” buttons on the toolbar in Prism .]

When starting a new block, I use the “Robot” button on the toolbar, select all translations, then click “Apply” to insert them in my translation.

When starting a new block, I use the “Robot” button on the toolbar, select some of the translations that I feel are usable, then click “Apply” to insert them in my translation.

As I work on a block, I use the “T” button on the toolbar to replace the segment I am currently working on with machine translation.

Other (please specify)

A.12. When importing machine translation output into a text you are working on, how much do you import at a time? (Select all that apply) (Optional question) [checkboxes]

A single segment

Multiple segments

A whole block

Other (please specify)

A.13. Have you encountered any problems with the Hawkeye user interface in Prism? (Optional question). [radio buttons]

No, I haven't

Yes, I have

A.14. Is there anything else you would like to say about the user interface? (Optional question). [text box]

B. Section for non-users of Hawkeye

B.1. Why do you not use the Hawkeye translation tool? (Optional question).

[text box]

Return to common section

4. Have you previously received training in using translation tools? If so, in what context? (Choose all that apply) (Optional question). [Matrix question]

	In a work environment	In an academic environment	Other	I have not received any training in the use of these tools
Machine translation tools and postediting				
Translation memory tools ⁴				
Terminology management tools ⁵				
Generative AI or other LLM-based tools ⁶				
Other translation tools				

⁴ **Translation memory tools** search previously translated texts for segments similar or identical to those in a text to be translated, and suggest the previous translations to the user for evaluation and possible use.

⁵ **Terminology management tools** store terms, phrases, or other recurrent items (usually along with their equivalents), and allow for manual searching or automatic retrieval of the stored items.

⁶ **Generative AI or Large Language Model (LLM)-based tools** are multi-purpose tools that function by analyzing large collections of data and using predictions based on this data to generate new material. They can be used to draft, translate, edit, correct, or adapt text, among many other tasks.

5. If you would like to provide more details about your training (e.g., what tools were covered, what you learned about them, in what other contexts you received training), you can do so here. (Optional question). [text box]

6. Do you currently use any other translation tools outside of Prism for your work in the Parliamentary service? (Optional question). [radio buttons]

No.

Yes.

[skip logic (No: continues to question 8, Yes: continues to question 7)]

7. What other translation tools do you use? (Optional question). [text box]

8. Have you ever worked in postediting? (Optional question). [radio buttons]

No

Yes, occasionally

Yes, as my main task

Yes, as my only task

Other (please specify)

[skip logic (No: continues to question 10, Yes: continues to question 9)]

9. How long did you work/have you worked in postediting (in any capacity)? (Optional question). [radio buttons]

Less than 6 months

6 months to less than 2 years

2 years to less than 5 years

5 years to less than 10 years

10 years or more

I am not sure

10. Which type of translation activity do you prefer? (Optional question). [radio buttons]

I would much rather translate from scratch than postedit

I would rather translate from scratch than postedit

I like translating from scratch and postediting equally

I would rather postedit than translate from scratch

I would much rather postedit than translate from scratch

I am not sure

11. Please describe your reason(s) for your preference given in the previous question (Optional question). [text box]

12. What tasks do you perform in your job? Choose all that apply. (Optional question). [checkboxes]

Translation: rendering of a written text in one natural language into another, whether entirely “from scratch” or assisted by one or more translation tools.

Revision: editing human translations, whether they were produced entirely “from scratch” or assisted by translation tools.

Postediting: editing the output of MT tools. This can be performed as a standalone task for full documents, or in the process of drafting parts of a translation.

Other (please specify)

13. In what language direction do you most frequently work? (Mandatory question). [radio buttons]

English to French

French to English

Equally French to English and English to French
Other (please specify)

14. How long have you been working in the language industry (e.g., as a translator or reviser), whether in your current workplace or elsewhere? (Optional question). [radio buttons]

Less than 2 years
2 to 5 years
6 to 10 years
11 to 15 years
16 to 20 years
More than 20 years
I am not sure

15. What kinds of documents do you work with? Check all that apply. (Optional question). [checkboxes]

House of Commons and Senate debates
House of Commons and Senate Committees

Parliamentary Documents
Other (please specify)

16. Is there anything that you have not been able to say above that you would like to add? Please feel free to add it here. (Optional question).

[text box]

That completes the survey. Thank you very much for your participation! Please note that after you click on the submit button, you will be automatically redirected to a separate (optional) form inviting you to participate in future interviews for this study.

[FINISH/SUBMIT BUTTON]

Thank you very much for your participation!

Before you leave...

To learn more about your experience with translation tools, including the machine translation tool integrated into the Prism system, we would like to interview some users. If you might be interested in participating and would like to know more, please enter your information in the form presented below. The information will remain confidential and will be stored separately from your questionnaire responses. We will contact you using the information you provide when we are ready to start recruitment.

If you do not wish to participate in the follow-up interviews, simply close this web page.

Interview sign-up form

Thank you for your interest in participating in our interviews! Please provide your name and the contact information you would like us to use to follow up with you. Please note that this information will be stored separately from your survey responses.

Name: [text box]

Email: [text box]

We will contact you with more details when we are ready to start recruiting for the interview phase of the study. You can decide if you would like to participate at that time.

Thank you again for your participation! We look forward to speaking with you.

Interview Protocol (Semi-structured)

Script

Introductory

Thank you for participating in this study.

In this interview, you will be asked questions regarding your experience with translation technologies, and specifically machine translation (MT) tools in your work. For that purpose, we will ask you about your experience and perception, the usefulness of machine translation tools for you, and how to improve their effectiveness to meet your needs.

The interview includes questions about the user interface, common issues with Machine Translation output quality and other technology needs you might have in your work. It also asks about your professional role and post-editing tasks. The interview should take about 45 minutes and it will be recorded.

As a reminder, your participation is completely voluntary and if you would like to revoke your consent you may do so at any time and we will end the interview and destroy the recording.

Do you have any questions for me before we begin?

Question Set (English)

Use of and Preferences for Other Tools (MT/CAT?)

1. Tell me about the *Computer-assisted translation* (CAT) tools that you use in your work for the parliamentary service.
2. If you could add any features to the CAT tools you use for your work at the Parliamentary Service, what would they be?
3. Have you ever used other Machine Translation tools, different from Hawkeye, for your work at the parliamentary service?
 - a. [Which ones?] (only if the answer to the previous question is YES)
 - b. [When you are using [tool or tools provided] instead of Hawkeye, what is something you can do that you could not do with Hawkeye?] (only if the answer to the previous question is YES)

MT Output Quality

4. Consider the types of errors you observe in the Hawkeye machine translation output. How do they compare against errors you might have observed in human translation?

Motivations for and Logistics of using the Hawkeye MT system

5. Do you work for both the Senate and the House of Commons?
6. [When working for the House of Commons, how often would you say you use the Hawkeye MT system?] (only if the answer to the previous question is not only Senate)

- a. [What factors would persuade you to use it more often?] (only if the answer is not ALL THE TIME)

User Interface

- 7. Imagine you're sitting at your computer, translating. Could you please explain the process you follow for using the Hawkeye Machine Translation system?
- 8. Does the user interface play a role in how you use the Hawkeye Machine Translation system?
 - a. [How so?] (only if the answer to the previous question is YES)

User Training

- 9. Could you please tell me about any training related to machine translation, post-editing, and the Hawkeye MT system you have received at your current workplace?
 - a. Do you think that receiving appropriate training affects your willingness to use the Hawkeye MT system?
 - b. What would you like to learn during training that you think would help you to use the machine translation system?

Role of translators vs. post-editors

- 10. Do you consider yourself a posteditor?
- 11. In terms of professional status, what is your perception of post-editing related to translation?
- 12. Do you think that post-editing influences your creative process of translation?
 - a. [How so?] (only if the answer to the previous question is YES)

Interview End

Do you have any questions for me?

- 1. Is there anything else you'd like to share or suggest?

Closing

We would like to thank you once again for your time and participation in our study. Your personal information will remain confidential and any personal identifying information will be removed from our records. Your recording will be destroyed once we have transcribed the interview. Is there anything else you would like to mention before we end this interview?

[...]

Thank you once again and have a nice day.