MMPE: A Multi-Modal Interface Using Handwriting, Touch Reordering, and Speech Commands for Post-Editing Machine Translation

Nico Herbig[1], Santanu Pal[1,2], Tim Düwel[1], Kalliopi Meladaki[1], Mahsa Monshizadeh[2], Vladislav Hnatovskiy[1], Antonio Krüger[1], Josef van Genabith[1,2]

[1]German Research Center for Artificial Intelligence (DFKI),
Saarland Informatics Campus, Germany
[2]Department of Language Science and Technology,
Saarland University, Germany
{firstname.lastname}@dfki.de
{firstname.lastname}@uni-saarland.de

Abstract

The shift from traditional translation to post-editing (PE) of machine-translated (MT) text can save time and reduce errors, but it also affects the design of translation interfaces, as the task changes from mainly generating text to correcting errors within otherwise helpful translation proposals. Since this paradigm shift offers potential for modalities other than mouse and keyboard, we present MMPE, the first prototype to combine traditional input modes with pen, touch, and speech modalities for PE of MT. Users can directly cross out or hand-write new text, drag and drop words for reordering, or use spoken commands to update the text in place. All text manipulations are logged in an easily interpretable format to simplify subsequent translation process research. The results of an evaluation with professional translators suggest that pen and touch interaction are suitable for deletion and reordering tasks, while speech and multi-modal combinations of select & speech are considered suitable for replacements and insertions. Overall, experiment participants were enthusiastic about the new modalities and saw them as useful extensions to mouse & keyboard, but not as a complete substitute.

1 Introduction & Related Work

As machine translation (MT) has been making substantial improvements in recent years[1], more and more professional translators are integrating this technology into their translation workflows (Zaretskaya et al., 2016; Zaretskaya and Seghiri, 2018). The process of using a pre-translated text as a basis and improving it to create the final translation is called post-editing (PE). While translation memory (TM) is still often valued higher than MT (Moorkens and O'Brien, 2017), a recent study by Vela et al. (2019) shows that professional translators chose PE of MT over PE of TM and translation from scratch in 80% of the cases. Regarding the time savings achieved through PE, Zampieri and Vela (2014) find that PE was on average 28% faster for technical translations, Toral et al. (2018) report productivity gains of 36% when using modern neural MT, and Aranberri et al. (2014) show that PE increases translation throughput for both professionals and lay users. Furthermore, it has been shown that PE not only leads to reduced time but also reduces errors (Green et al., 2013).

Switching from traditional translation to PE results in major changes in translation workflows (Zaretskaya and Seghiri, 2018), including the interaction pattern (Carl et al., 2010), yielding a significantly reduced amount of mouse and keyboard events (Green et al., 2013). This requires thorough investigation in terms of interface design, since the task changes from mostly text production to comparing and adapting MT and TM proposals, or put differently, from control to supervision.

While most computer-aided translation (CAT) tools focus on traditional translation and incorporate only mouse & keyboard, previous research investigated other input modalities: automatic speech recognition (ASR) for dictating translations has already been explored in the 90s (Dymetman et al., 1994; Brousseau et al., 1995) and the more recent investigation of ASR for PE (Martinez et al., 2014) even argues that a combination with typing could boost productivity. Mesa-Lao (2014) finds that PE trainees have a positive attitude towards speech input and would consider adopting it, and Zapata et al. (2017) found that ASR for PE was faster than ASR for translation from scratch. Due to these benefits, commercial CAT tools like memoQ and MateCat are also beginning to integrate ASR.

The CASMACAT tool (Alabau et al., 2013) allows the user to input text by writing with e-pens in

[1] WMT 2019 translation task: http://matrix.statmt.org/, accessed 07. Jan 2020

a special area. A vision paper (Alabau and Casacuberta, 2012) proposes to instead use e-pens for PE sentences with few errors in place and provides examples of symbols that could be used for this. Studies on mobile PE via touch and speech (O'Brien et al., 2014; Torres-Hostench et al., 2017) show that participants especially liked reordering words through touch drag and drop, and preferred voice when translating from scratch, but used the iPhone keyboard for small changes. Teixeira et al. (2019) also explore a combination of touch and speech; however, their touch input received poor feedback since (a) their tile view (where each word is a tile that can be dragged around) made reading more complicated, and (b) touch insertions were rather complex to achieve within their implementation. In contrast, dictation functionality was shown to be quite good and even preferred to mouse and keyboard by half of the participants. The results of an elicitation study by Herbig et al. (2019a) indicate that pen, touch, and speech interaction should be combined with mouse and keyboard to improve PE of MT. In contrast, other modalities like eye tracking or gestures were seen as less promising.

This paper presents MMPE, the first translation environment combining standard mouse & keyboard input with touch, pen, and speech interactions for PE of MT. It allows users to directly cross out or hand-write new text, drag and drop words for reordering, or use spoken commands to update the text in place. All text manipulations are logged in an easily interpretable format (e.g., *replaceWord* with the old and new word) to facilitate translation process research. The results of a study with 11 professional translators show that participants are enthusiastic about having these alternatives, and suggest that pen and touch are well suited for deletion and reordering operations, whereas speech and multi-modal interaction are suitable for insertions and replacements.

2 The MMPE Prototype

This section presents the MMPE prototype (see Figure 1), which combines pen, touch, and speech input with a traditional mouse and keyboard approach for PE of MT. The prototype is designed for professional translators in an office setting. A video demonstration is available at `https://youtu.be/tkJ9OWmDd0s`.

2.1 Apparatus

On the software side, we decided to use Angular[2] for the frontend, and node.js[3] for the backend.

The frontend, including all of the newly implemented modalities for text editing, is what the system currently focuses on. While this Angular frontend could be used in a browser on any device, we initially design for the following hardware to optimally support the implemented interactions: we use a large tiltable touch & pen screen (see Figure 1a), namely the Wacom Cintiq Pro 32 inch display. Together with the Flex Arm, this screen can be moved up in the air to work in a standing position, or it can be tilted and moved flat on the table (similar to how users use a tablet), thereby supporting better pen and touch interaction (as requested in Herbig et al. (2019a)). To avoid limitations in ASR through a potentially bad microphone, we further use the Sennheiser PC 8 Headset for speech input. Last, mouse and keyboard are provided.

Since it is not the focus of this work, the backend is kept rather minimal: it allows saving and loading of projects (including the MT) from JSON files, can store log files, etc. Here, the project files simply contain an array of segments with source, target, as well as any MT or TM proposal that should initially be shown for PE.

2.2 Overall Layout

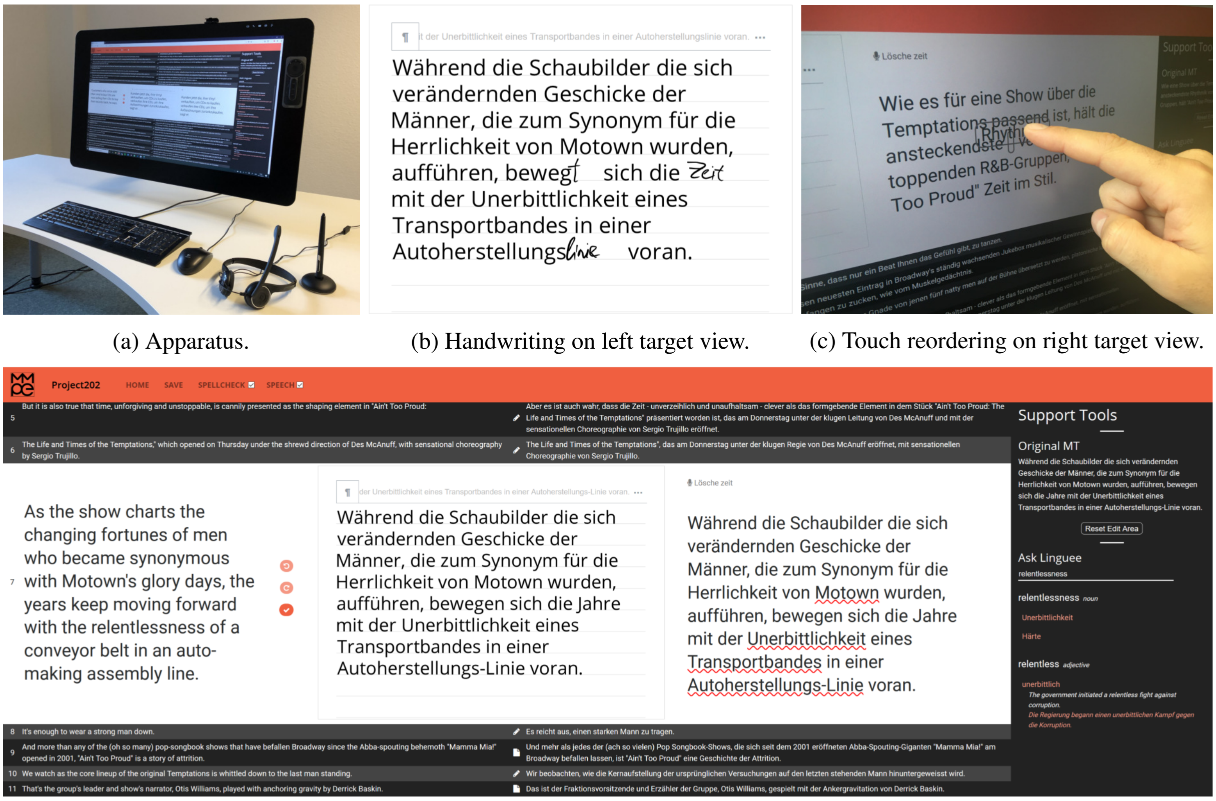
Figure 1d shows our implemented horizontal source-target layout, where each segment's status (unedited, edited, confirmed) is visualized between source and target. On the far right, support tools are offered as requested in Herbig et al. (2019a): (1) the unedited MT output, to which the user can revert his editing using a button, and (2) a corpus combined with a dictionary: when entering a word or clicking/touching a word in the source view on the left, the Linguee[4] website is queried to show the word in context and display its primary and alternative translations. The top of the interface shows a toolbar where users can enable or disable speech recognition as well as spell checking, save and load projects, or navigate to another project.

The current segment is enlarged, thereby offering space for handwritten input and allowing the user to view a lot of context while still seeing the current segment in a comfortable manner (Herbig

[2] https://angular.io/, accessed 07. Jan 2020

[3] https://nodejs.org/en/, accessed 07. Jan 2020

[4] https://www.linguee.com/, accessed 07. Jan 2020



¶ t der Unerbittlichkeit eines Transportbandes in einer Autoherstellungslinie voran. ...

Während die Schaubilder die sich verändernden Geschicke der Männer, die zum Synonym für die Herrlichkeit von Motown wurden, aufführen, bewegt sich die Zeit mit der Unerbittlichkeit eines Transportbandes in einer Autoherstellungslinie voran.

(a) Apparatus.

(b) Handwriting on left target view.

(c) Touch reordering on right target view.

5 But it is also true that time, unforgiving and unstoppable, is cannily presented as the shaping element in "Ain't Too Proud: — Aber es ist auch wahr, dass die Zeit - unverzeihlich und unaufhaltsam - clever als das formgebende Element in dem Stück "Ain't Too Proud: The Life and Times of the Temptations" präsentiert worden ist, das am Donnerstag unter der klugen Leitung von Des McAnuff und mit der sensationellen Choreographie von Sergio Trujillo eröffnet.

6 The Life and Times of the Temptations," which opened on Thursday under the shrewd direction of Des McAnuff, with sensational choreography by Sergio Trujillo. — The Life and Times of the Temptations", das am Donnerstag unter der klugen Regie von Des McAnuff eröffnet, mit sensationellen Choreographie von Sergio Trujillo.

7 As the show charts the changing fortunes of men who became synonymous with Motown's glory days, the years keep moving forward with the relentlessness of a conveyor belt in an auto-making assembly line.

¶ der Unerbittlichkeit eines Transportbandes in einer Autoherstellungs-Linie voran. ...

Während die Schaubilder die sich verändernden Geschicke der Männer, die zum Synonym für die Herrlichkeit von Motown wurden, aufführen, bewegen sich die Jahre mit der Unerbittlichkeit eines Transportbandes in einer Autoherstellungs-Linie voran.

Lösche zeit

Während die Schaubilder die sich verändernden Geschicke der Männer, die zum Synonym für die Herrlichkeit von Motown wurden, aufführen, bewegen sich die Jahre mit der Unerbittlichkeit eines Transportbandes in einer Autoherstellungs-Linie voran.

8 It's enough to wear a strong man down. — Es reicht aus, einen starken Mann zu tragen.

9 And more than any of the (oh so many) pop-songbook shows that have befallen Broadway since the Abba-spouting behemoth "Mamma Mia!" opened in 2001, "Ain't Too Proud" is a story of attrition. — Und mehr als jedes der (ach so vielen) Pop Songbook-Shows, die sich seit dem 2001 eröffneten Abba-Spouting-Giganten "Mamma Mia!" am Broadway befallen lassen, ist "Ain't Too Proud" eine Geschichte der Attrition.

10 We watch as the core lineup of the original Temptations is whittled down to the last man standing. — Wir beobachten, wie die Kernaufstellung der ursprünglichen Versuchungen auf den letzten stehenden Mann hinuntergeweisst wird.

11 That's the group's leader and show's narrator, Otis Williams, played with anchoring gravity by Derrick Baskin. — Das ist der Fraktionsvorsitzende und Erzähler der Gruppe, Otis Williams, gespielt mit der Ankergravitation von Derrick Baskin.

Support Tools

Original MT

Während die Schaubilder die sich verändernden Geschicke der Männer, die zum Synonym für die Herrlichkeit von Motown wurden, aufführen, bewegen sich die Jahre mit der Unerbittlichkeit eines Transportbandes in einer Autoherstellungs-Linie voran.

Reset Edit Area

Ask Linguee

relentlessness

relentlessness *noun*

Unerbittlichkeit

Härte

relentless *adjective*

unerbittlich

The government initiated a relentless fight against corruption.

Die Regierung begann einen unerbittlichen Kampf gegen die Korruption.

(d) Screenshot of the interface.

Figure 1: Overview of the MMPE prototype.

et al. (2019a)). The view for the current segment is further divided into the source segment (left) and two editing planes for the target, one for handwriting and drawing gestures (middle), and one for touch deletion & reordering, as well as standard mouse and keyboard input (right). Both initially show the MT proposal, and synchronize on changes to either one. The reason for having two editing fields instead of only one is that some interactions are overloaded, e.g., a touch drag can be interpreted as both hand-writing (middle) and reordering (right). Undo and redo functionality for all modalities, as well as confirming segments, are also implemented through buttons between the source and target texts, and can further be triggered through hotkeys. The target text is spell-checked, as a lack of this feature was criticized in Teixeira et al. (2019).

2.3 Left Target View: Handwriting

For handwriting recognition (see Figure 1b), we use the MyScript Interactive Ink SDK[5]. Apart from merely recognizing the written input, it offers gestures[6] like strike-through or scribble for deletions, breaking a word into two (draw line from top to bottom), and joining words (draw line from bottom to top). For inserting words, one can directly write into empty space, or create such space first by breaking the line (draw a long line from top to bottom), and hand-writing the word then. All changes are immediately interpreted, i.e., striking through a word deletes it immediately instead of showing it in a struck-through visualization. While it is not necessary to convert text from the hand-written appearance into computer font, the user can do so using a small button at the top of the editor. The editor further shows the recognized handwritten text immediately at the very top of the drawing view in a small gray font, where alternatives for the current recognition are offered when clicking on a recognized word. Since all changes from this drawing view are immediately synchronized into the right-hand view, the user can also see the recognized text there. Apart from using the pen, the user can use his/her finger or the mouse on the left-hand editing view for hand-writing.

[5] https://developer.myscript.com/, accessed 07. Jan 2020

[6] https://developer.myscript.com/docs/concepts/editing-gestures/, accessed 07. Jan 2020

2.4 Right Target View: Touch Reordering, Mouse & Keyboard

On the right-hand editing view, the user can delete words by simply double-tapping them with pen/finger touch, or reorder them through a simple drag and drop procedure (see Figure 1c). This procedure visualizes the picked-up word as well as the current drop position through a placeholder element. Spaces between words and punctuation marks are automatically fixed, i.e., double spaces at the pickup position and missing spaces at the drop position are corrected. This reordering functionality is strongly related to Teixeira et al. (2019); however, only the currently dragged word is temporarily visualized as a tile to offer better readability. Furthermore, the cursor can be placed between words using a single tap, allowing the user to combine touch input with e.g., the speech or keyboard modalities (see below). Naturally, the user can also edit and navigate using mouse and keyboard, where all common shortcuts work as expected from other software (e.g., ctrl+arrow keys or ctrl+c).

2.5 Speech Input

To minimize lag during speech recognition, we use a streaming approach, sending the recorded audio to IBM Watson servers to receive a transcription, which is then interpreted in a command-based fashion. Thus, our speech module not only handles dictations as in Teixeira et al. (2019) but can correct mistakes in place.

The transcription itself is visualized at the top of the right target view (see Figure 1c). As commands, the user has the option to "*insert*", "*delete*", "*replace*", and "*reorder*" words or subphrases. To specify the position if it is ambiguous, one can define anchors as in "*after*"/"*before*"/"*between*", or define the occurrence of the token ("*first*"/"*second*"/"*last*"). A full example is "*insert A after second B*", where A and B can be words or subphrases. In contrast to the other modalities, character-level commands are not supported, so instead of deleting an ending, one should replace the word. Again, spaces between words and punctuation marks are automatically fixed upon changes. For the German language, nouns are automatically capitalized using the list of nouns from Wiktionary[7].

2.6 Multi-modal Combinations

Last, the user can use a multi-modal combination, i.e., pen/touch/mouse combined with speech. For this, a target word/position first needs to be specified by placing the cursor on or next to a word using the pen, finger touch, or the mouse/keyboard; alternatively, the word can be long-pressed with pen/touch. Afterwards, the user can use a voice command like *"delete"*, *"insert A"*, *"move after/before A/between A and B"*, or *"replace by A"* without needing to specify the position/word, thereby making the commands less complex.

2.7 Logging

We implemented extensive logging functionality: on the one hand, we log the concrete keystrokes, touched pixel coordinates, etc.; on the other hand, all UI interactions (like *segmentChange* or *undo/redo/confirm*) are stored, allowing us to analyze the translator's use of MMPE.

Most importantly, however, we also log all text manipulations on a higher level to simplify text editing analysis: for *insertions*, we log whether a single or multiple words were inserted, and add the actual words and their positions as well as the segment's content before and after the insertion to the log entry. *Deletions* are logged analogously, and for *reorderings*, we add the old and the new position of the moved words to the log entry. Last, for *replacements*, we log whether only a part of a word was replaced (i.e., changing the word form), whether the whole word was replaced (i.e., correcting the lexical choice), or whether a group of words was replaced. In all cases, the words before and after the change, as well as their positions and the overall segment text are specified in the log entry.

Furthermore, all log entries contain the modality that was used for the interaction, e.g., Speech or Pen, thereby allowing the analysis of which modality was used for which editing operation. All log entries with their timestamps are created within the Angular client and sent to the node.js server for storage in a JSON file.

3 Evaluation

We evaluated the prototype with 11 professional translators[8]. Since our participants were German

[7] https://en.wiktionary.org/wiki/Category:German_noun_forms, accessed 07. Jan 2020

[8] The study has been approved by the university's ethical review board, and participants were paid for their time. The data and analysis scripts can be found at `https://mmpe.dfki.de/data/ACL2020/`.

natives, we chose a EN-DE translation task to avoid ASR recognition errors occurring in non-native commands (Dragsted et al., 2011). In the following, "modalities" refers to Touch (T), Pen (P), Speech (S), Mouse & Keyboard (MK), and Multi-Modal combinations (MM, see Section 2.6), while "operations" refers to Insertions, Deletions, Replacements, and Reorderings. More details on the evaluation are presented in Herbig et al. (2020).

3.1 Method

The study took approximately 2 hours per participant and involved three separate stages. First, participants filled in a questionnaire capturing demographics as well as information on CAT usage. In stage two, participants received an explanation of all of the prototype's features and then had 10–15 minutes to explore the prototype on their own and become familiar with the interface. Finally, stage three included the main experiment, which is a guided test of all implemented features combined with Likert scales and interviews, as described in detail below.

The main part tests each of the 5 modalities for each of our 4 operations in a structured way. For this, we prepared four sentences for each operation by manually introducing errors into the reference sentences from the WMT news test set 2018. Thus, overall each participant had to correct 4 segments per operation (4) using each modality (5), which results in $4 \times 4 \times 5 = 80$ segments. Within the four sentences per operation, we tried to capture slightly different cases, like deleting single words or a group of words. The prototype was adapted for this controlled task such that it displays a popup when selecting a segment, visualizing the necessary correction to apply as well as the modality to use. The reason why we provided the correction to apply was to ensure a consistent editing behavior across all participants, thereby making the following measurements comparable: each modality had to be rated for each operation on 7-point Likert scales assessing whether the modality is a good fit, whether it is easy to use, and whether it is a good alternative to MK. Furthermore, participants had to order the modalities from best to worst for each operation. Last, we captured their comments in an interview after each operation and measured the times required to fix the introduced errors. In the end, a final unstructured interview to capture high-level feedback on the interface was conducted.

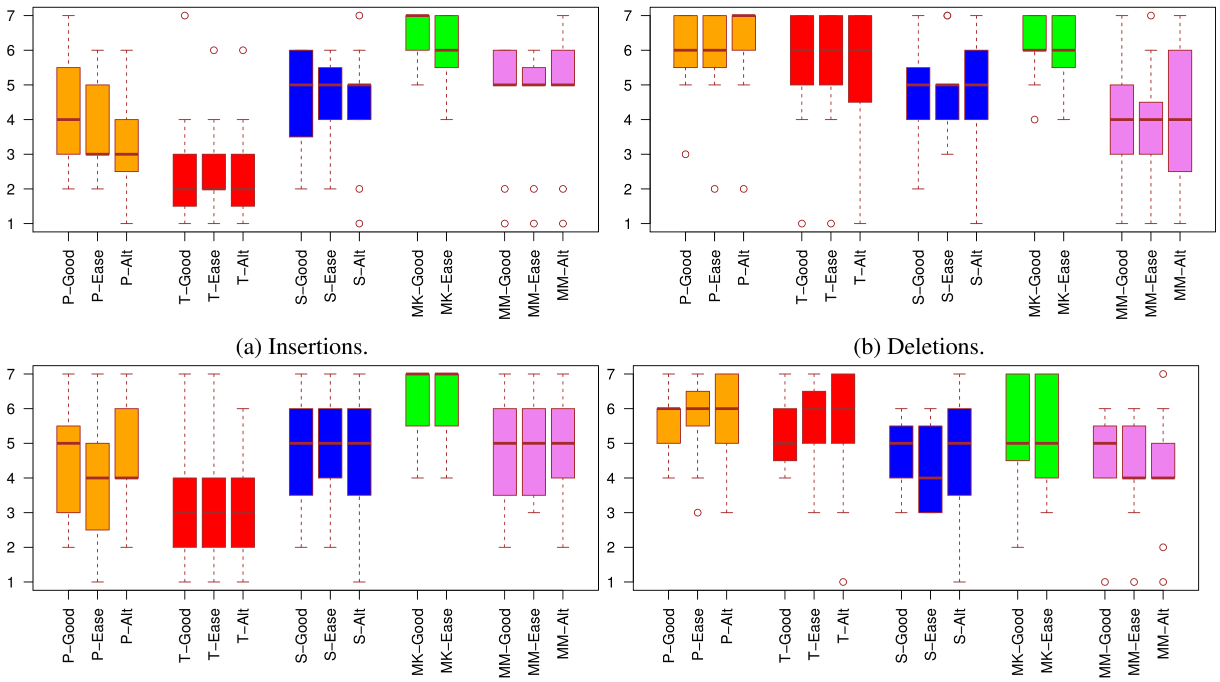
3.2 Results & Discussion

Figure 2 depicts the results of the 3 Likert scales of the 5 modalities for the 4 tasks. The participants' orderings of modalities for the operations were mostly in line with these ratings, as we will discuss in the next sections.

According to subjective ratings, modality ordering, and comments, ***P(en)*** is among the best modalities for deletions and reordering. However, other modalities are superior for insertions and replacements, where P was seen as suitable only for short modifications, and to be avoided for more extended changes. In terms of timings, P was also among the fastest for deletions and reorderings, and among the slowest for insertions. What is interesting, however, is that P was significantly faster than S and MM for replacements (by 6 and 7 seconds on average) even though it was rated lower. Participants also commented very enthusiastically about pen reordering and deletions, as they would nicely resemble manual copy-editing. The main concern for hand-writing was the need to think about and to create space before actually writing.

Results for ***T(ouch)*** were similarly good for deletions and reorderings, but it was considered worse for insertions and replacements. Furthermore, and as we expected due to its precision, pen was preferred to finger touch by most participants. However, in terms of timings, the two did not differ significantly, apart from replace operations (where pen was faster). Even for replacements, where T was rated as the worst modality, it actually was (non-significantly) faster than S and MM.

S(peech) and ***M(ulti)-M(odal)*** PE were considered the worst and were also the slowest modalities for reordering and deletions. For insertions and replacements, however, these two modalities were rated and ordered 2$^{\text{nd}}$ (after MK) and in particular much better than P and T. Timing analysis agrees for insertions, being 2$^{\text{nd}}$ after MK; for replacements, however, S and MM were the slowest even though the ratings put them ahead of P and T. Insertions are the only operation where MM was (non-significantly) faster than S, since the position did not have to be verbally specified. Even though participants were concerned regarding formulating commands while mentally processing text, they considered S and MM especially interesting for adding longer text. The main advantage of MM would be that one has to speak less, albeit at the cost of doing two things at once.



(a) Insertions.

(b) Deletions.

(c) Replacements.

(d) Reorderings.

Figure 2: Subjective ratings of the five modalities for the four operations on the 7-point Likert scales for goodness, ease of use, and whether it is a good alternative to MK.

M(ouse) & K(eyboard) received the best scores for insertions and replacements, where it was also the fastest. Furthermore, it got good ratings for deletions and reorderings. For deletions, MK was comparably fast to P, T, and S. For reordering, however, it was slower than P and T. Some participants commented negatively on MK, stating that it only works well because of "years of expertise", and being "unintuitive" especially for reordering.

Overall, many participants provided very positive feedback on this first prototype combining pen, touch, speech, and multi-modal combinations for PE MT, encouraging us to continue. They especially highlighted that it was nice to have the option to switch between modalities. Furthermore, several promising ideas for improving the prototype were proposed, e.g., to visualize whitespaces.

4 Conclusion

While more and more professional translators are switching to the use of PE to increase productivity and reduce errors, current CAT interfaces still heavily focus on traditional mouse and keyboard input. This paper therefore presents MMPE, a CAT prototype combining pen, touch, speech, and multi-modal interaction together with common mouse and keyboard input possibilities. Users can directly cross out or hand-write new text, drag and drop words for reordering, or use spoken commands to update the text in place. Our study with professional translators shows a high level of interest and enthusiasm about using these new modalities. For deletions and reorderings, pen and touch both received high subjective ratings, with pen being even better than mouse & keyboard. For insertions and replacements, speech and multi-modal interaction were seen as suitable interaction modes; however, mouse & keyboard were still favored and faster.

As a next step, we will improve the prototype based on the participants' valuable feedback. Furthermore, an eye tracker will be integrated into the prototype that can be used in combination with speech for cursor placement, thereby simplifying multi-modal PE. Last, we will investigate whether using the different modalities has an impact on cognitive load during PE (Herbig et al., 2019b).

Acknowledgments

This research was funded in part by the German Research Foundation (DFG) under grant number GE 2819/2-1 (project MMPE). We thank AMPLEXOR (`https://www.amplexor.com`) for their excellent support in providing access to professional human translators for our experiments.

References

Vicent Alabau, Ragnar Bonk, Christian Buck, Michael Carl, Francisco Casacuberta, Mercedes García-Martínez, Jesús González, Philipp Koehn, Luis Leiva, Bartolomé Mesa-Lao, et al. 2013. CASMACAT: An open source workbench for advanced computer aided translation. *The Prague Bulletin of Mathematical Linguistics*, 100:101–112.

Vicent Alabau and Francisco Casacuberta. 2012. Study of electronic pen commands for interactive-predictive machine translation. In *Proceedings of the International Workshop on Expertise in Translation and Post-Editing – Research and Application, Copenhagen, Denmark*, pages 17–18.

Nora Aranberri, Gorka Labaka, A Diaz de Ilarraza, and Kepa Sarasola. 2014. Comparison of post-editing productivity between professional translators and lay users. In *Proceeding of AMTA Third Workshop on Post-editing Technology and Practice*, pages 20–33.

Julie Brousseau, Caroline Drouin, George Foster, Pierre Isabelle, Roland Kuhn, Yves Normandin, and Pierre Plamondon. 1995. French speech recognition in an automatic dictation system for translators: The TransTalk project. In *Proceedings of Eurospeech Fourth European Conference on Speech Communication and Technology*, pages 193–196.

Michael Carl, Martin Jensen, and Kay Kristian. 2010. Long distance revisions in drafting and post-editing. *CICLing Special Issue on Natural Language Processing and its Applications*, pages 193–204.

Barbara Dragsted, Inger Margrethe Mees, and Inge Gorm Hansen. 2011. Speaking your translation: Students' first encounter with speech recognition technology. *Translation & Interpreting*, 3(1):10–43.

Marc Dymetman, Julie Brousseau, George Foster, Pierre Isabelle, Yves Normandin, and Pierre Plamondon. 1994. Towards an automatic dictation system for translators: The TransTalk project. In *Proceedings of the ICSLP International Conference on Spoken Language Processing*.

Spence Green, Jeffrey Heer, and Christopher D Manning. 2013. The efficacy of human post-editing for language translation. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 439–448. ACM.

Nico Herbig, Tim Düwel, Santanu Pal, Kalliopi Meladaki, Mahsa Monshizadeh, Antonio Krüger, and Josef van Genabith. 2020. MMPE: A multi-modal interface for post-editing machine translation. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. Association for Computational Linguistics.

Nico Herbig, Santanu Pal, Josef van Genabith, and Antonio Krüger. 2019a. Multi-modal approaches for post-editing machine translation. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, page 231. ACM.

Nico Herbig, Santanu Pal, Mihaela Vela, Antonio Krüger, and Josef Genabith. 2019b. Multi-modal indicators for estimating perceived cognitive load in post-editing of machine translation. *Machine Translation*, 33(1-2):91–115.

Mercedes Garcia Martinez, Karan Singla, Aniruddha Tammewar, Bartolomé Mesa-Lao, Ankita Thakur, MA Anusuya, Banglore Srinivas, and Michael Carl. 2014. SEECAT: ASR & eye-tracking enabled computer assisted translation. In *The 17th Annual Conference of the European Association for Machine Translation*, pages 81–88. European Association for Machine Translation.

Bartolomé Mesa-Lao. 2014. Speech-enabled computer-aided translation: A satisfaction survey with post-editor trainees. In *Proceedings of the EACL 2014 Workshop on Humans and Computer-assisted Translation*, pages 99–103.

Joss Moorkens and Sharon O'Brien. 2017. Assessing user interface needs of post-editors of machine translation. In *Human Issues in Translation Technology*, pages 127–148. Routledge.

Sharon O'Brien, Joss Moorkens, and Joris Vreeke. 2014. Kanjingo – a mobile app for post-editing. In *Proceedings of the 17th Annual Conference of the European Association for Machine Translation*.

Carlos S.C. Teixeira, Joss Moorkens, Daniel Turner, Joris Vreeke, and Andy Way. 2019. Creating a multimodal translation tool and testing machine translation integration using touch and voice. *Informatics*, 6.

Antonio Toral, Martijn Wieling, and Andy Way. 2018. Post-editing effort of a novel with statistical and neural machine translation. *Frontiers in Digital Humanities*, 5:9.

Olga Torres-Hostench, Joss Moorkens, Sharon O'Brien, Joris Vreeke, et al. 2017. Testing interaction with a mobile MT post-editing app. *Translation & Interpreting*, 9(2):138.

Mihaela Vela, Santanu Pal, Marcos Zampieri, Sudip Kumar Naskar, and Josef van Genabith. 2019. Improving CAT tools in the translation workflow: New approaches and evaluation. In *Proceedings of Machine Translation Summit XVII Volume 2: Translator, Project and User Tracks*, pages 8–15.

Marcos Zampieri and Mihaela Vela. 2014. Quantifying the influence of MT output in the translators' performance: A case study in technical translation. In *Proceedings of the EACL 2014 Workshop on Humans and Computer-assisted Translation*, pages 93–98.

Julián Zapata, Sheila Castilho, and Joss Moorkens. 2017. Translation dictation vs. post-editing with cloud-based voice recognition: A pilot experiment. *Proceedings of MT Summit XVI*, 2.

Anna Zaretskaya and Míriam Seghiri. 2018. *User Perspective on Translation Tools: Findings of a User Survey*. Ph.D. thesis, University of Malaga.

Anna Zaretskaya, Mihaela Vela, Gloria Corpas Pastor, and Miriam Seghiri. 2016. Comparing post-editing difficulty of different machine translation errors in Spanish and German translations from English. *International Journal of Language and Linguistics*, 3(3):91–100.