



Chapitre d'actes

2023

Published version

Open Access

This is the published version of the publication, made available in accordance with the publisher's policy.

A pedagogical platform for spoken post-editing (PE): the integration of speech input into COPECO

Liyana Pathirana, Jeevanthi; Bouillon, Pierrette; Mutal, Jonathan David

How to cite

LIYANA PATHIRANA, Jeevanthi, BOUILLON, Pierrette, MUTAL, Jonathan David. A pedagogical platform for spoken post-editing (PE): the integration of speech input into COPECO. In: International Conference on Human-Informed Translation and Interpreting Technology (HiT-IT 2023). Constantin Orăsan, Ruslan Mitkov, Gloria Corpas Pastor and Johanna Monti (Ed.). Naples, Italy. [s.l.] : [s.n.], 2023. p. 195–202. doi: 10.26615/issn.2683-0078.2023_018

This publication URL: <https://archive-ouverte.unige.ch/unige:170861>

Publication DOI: [10.26615/issn.2683-0078.2023_018](https://doi.org/10.26615/issn.2683-0078.2023_018)

© The author(s). This work is licensed under a Creative Commons Attribution (CC BY 4.0)

<https://creativecommons.org/licenses/by/4.0>

A pedagogical platform for spoken post-editing (PE): the integration of speech input into COPECO

Jeevanthi Liyana Pathirana¹, Pierrette Bouillon² and Jonathan Mutal³

¹²³ Faculty of Translation and Interpreting, University of Geneva
juliyana@pathirana@gmail.com, Pierrette.Bouillon@unige.ch,
Jonathan.Mutal@unige.ch

Abstract. We present the integration of speech input into COPECO, an online teaching and learning platform developed to collect learner translation/post-editing corpora[22]. Speech offers potential to increase translators' productivity and wellbeing by reducing typing time and effort. A multimodal, speech-enabled COPECO will complement other researchers' and developers' efforts to integrate speech dictation and post-editing into translation tools, training and practice (e.g. MateCat allows post-editors to activate a speech-to-text component or TradDICT Learn[28], an online-learning platform to develop sight-translation and dictation skills). COPECO helps trainers to compare different translation modalities, prepare courses and share teaching resources. To our knowledge, this is the first online platform to allow comparison between written and spoken post-editing and gathering data/statistics on speech-based post-editing behaviour (e.g. respeaking whole segments vs. correcting minor errors, etc). This helps translation trainers understand frequent errors made by learners when using speech which can then be used to improve their course content. It will also help to build a "speech post-editing corpus" which will be beneficial in the long run for research and analytical purposes. Unlike proprietary speech recognition add-ons with defined services, we develop speech commands for post-editing based on translators' behaviour and needs, optimising COPECO based on requirements with minimal costs.

Keywords: Automatic Speech Recognition (ASR), Post editing (PE), Computer Assisted Translation (CAT)

1 Introduction

Translation services are used for translating documents into multiple languages, within a limited time frame, with high accuracy. Post-editing of Machine Translation (MT) is known to allow translating large volumes of translations while saving costs and time [29]. Workflows in the translation industry have experienced a significant transformation in a way that speech technology is likely to contribute to further innovation [8]. Preliminary studies on speech based post-editing [14][15][16] show that provided that Automatic Speech Recognition (ASR) and MT output are of high quality and that the translators are competent with software (Computer Assisted Transla-

tion (CAT) tools, MT suggestions and ASR toolkits such as Dragon [9]), speech based post-editing can be a promising approach which can result in performance gains in the translation workflow. However, for speech based post-editing to be used by translators, this modality has to be properly introduced to them. For that, translation trainers should analyze behaviors of speech based post-editing and translation, to then use that knowledge to improve their courses on translation training.

In this paper we demonstrate COPECO [22], a speech-enabled platform developed for collecting learner translation/post-editing corpora and for helping translation teachers annotate and learn student post-editing behavior. First we give an overview of speech based translation/post-editing in previous studies, leading towards the rationale behind our work and then we demonstrate the design of our setup and its functionalities.

2 Speech-based PE

Computer Assisted Translation tools are mostly based on traditional input such as keyboard and mouse [27]. However, the translation industry constantly seeks ways to improve speed and quality by incorporating various technological advancements into translation workflows [3][21][24].

One such technology is Automatic Speech Recognition (ASR) which automatically transcribes spoken input into text[7]. As a result, a large body of literature has explored the integration of ASR into translation processes and the ways in which speech technologies can be effectively utilized in translation[8][10][20][25][27][32][34]. Previous studies described successful use of ASR by freelance translators, with more productivity, allowing for "more flexible, translator-centered, ergonomic workflows and workspaces"[6][7]. Commercially available CAT tools have started offering integration with ASR systems as well, aiming for increased productivity and ergonomics during the translation process. Some examples are memoQ [19] combined with Apple's speech recognition service or Matecat combined with the ability to dictate the translation[17][18]. With some other CAT tools [30], commercial ASR systems for dictation, such as Dragon Naturally Speaking[9] can be used.

Recent studies researched and surveyed [20] the potential of using ASR for post-editing purposes. A study investigating the effects on productivity and on a translator's experience of integrating machine translation post-editing with speech technologies revealed that post-editing with the aid of a speech recognition system was faster than translating with the aid of a speech recognition system and also less tiresome (i.e., more ergonomic)[34]. Similarly, studies that looked into the possibility of using speech technologies for post-editing purposes in international organizations revealed translators were open to try speech-based post-editing as a new translation workflow [14]. Another study[11] found voice input more interesting than typing only when post-editing, as when some segments require major changes they could be dictated. If the post-editor is not a touch-typist, it was also found that the back and forth transfer of visual attention between source text, machine translation output and keyboard adds to the complexity of the task.

Many studies investigated the possibility of having multimodal interfaces to improve translation workflows, where ASR was a component. One study observes and analyzes translator experience (TX) with off-the-shelf voice-and-touch-enabled multimodal interfaces, as opposed to the interaction with traditional keyboard-and-mouse graphical user interfaces, to provide better recommendations for translation tool design [33]. Later, [27] developed a web-based translation editing interface that permits multimodal input via touch-enabled screens and speech recognition in addition to keyboard and mouse, which demonstrated the importance of the ASR quality rather than the features of the interface. [12] presented and evaluated the MMPE CAT environment that explores the use of speech commands, handwriting input, touch reordering, and multi-modal combinations for PE of MT. MMPE was later improved with additional speech and other facilities in the next version [13] along with some correction commands based on speech input. Unlike [27] which used a web-based interface, this environment requires specific hardware, which obliges the users to have this configuration in their translation environment.

These studies confirm that speech as an input modality does indeed provide promising improvements in the translation workflow, which supports it being included in the translator training process. Previous studies[2] mention current occurrences of audiovisual translation (AVT) modules incorporated in translator training programs, where use of ASR and TTS in AVT is specially studied in specialist courses. The study also emphasizes the importance of translator trainers being acquainted with the professional environment and the latest trends. There already exist courses such as Tradict Learn [28] which provides introductory courses to learn interactive translation (using voice) as well.

However, if we want to understand the frequent errors and behaviors of these modalities, it is necessary to analyze and understand translator data and statistics. Today, many tools exist for monitoring written-based PE (e.g. PET [1]), but when it comes to analyzing speech based inputs for translator training, there are only a few examples. Workbenches like Matecat [17] enable speech based input for translating via dictation [18], but do not contain enough data protection or the ability to revise using customized annotation schemes. A recent study experiments with the combination of speech synthesis and PE where they investigate the benefits and drawbacks of exposing students to novel technologies and practices such as synthetic voices and PE early, and the role these interactions can play in translator training [4]. However, very few have considered using data for inferring translator training techniques on ASR based techniques.

3 COPECO Design

Most previous studies on speech based post-editing and translation have been based on tools that require specific hardware requirements based standalone applications and other studies require commercial license based software such as Dragon[8] and Trados Studio workbench [30]. For our work, we chose the tailor-made open source PE platform COPECO [22] to integrate speech recognition. COPECO was originally

developed as a project to collect post-edits produced by students and teacher corrections and to structure the task of translation error annotation. The aim of COPECO is to translation teachers with an online post-editing platform, designed to help them to annotate student post-editing tasks using a shared or personalized annotation scheme. In the long run, data collected from a platform like COPECO can also be used to analyze how post-editing tools can impact the translation industry economy.

These features made COPECO a suitable platform to integrate speech so that speech based translation or post-editing can be done. Currently, speech integrated COPECO web based platform allows the translation trainers to assign tasks with text to students to translate and their machine translation suggestions. The student can then translate from scratch or post-edit the translation, using either typing, speech, or a mix of both. Once the task is translated, the student can submit the task to the teacher. The teacher can then correct the task using systematic translation error annotation [23]. Previous studies have worked on defining frameworks on error analysis of MT [31] or ASR[26] outputs. COPECO allows error annotation with predefined translation schemes or their custom annotation schemes. COPECO also allows to visualize the corpus with the translations, corrections, reference translation (if it exists) as well as the annotations. It simultaneously builds an open source student post-editing corpus by collecting post-edits produced by student and teacher corrections. All data are collected and can be anonymously shared.

For speech integration, we used a publicly available ASR engine powered by Google Web Speech, which connected to our tool via an application programming interface (API)[5]. The Web speech API only allows speech recognition, and no commands were available for post-editing. So an initial set of speech commands used for post-editing were developed. Some of the commands developed for English include "clear segment", moving the cursor to specific locations, selecting a word/phrase and replacing it, deleting words/phrases, saving a segment and moving to the next segment. Table 01 shows a subset of the speech commands that works for English language. Similarly, speech commands can be developed for any other target language that we would need to use speech based translation and post-editing tasks.

Table 1. Example set of speech commands for post-editing in English

Command	Description
"ABC"	Inserts ABC in the current cursor location.
"Select XYZ"	Highlights XYZ (can be a word or a phrase).
"Undo Highlight"	Removes the highlighting of XYZ.
"Delete that"	Deletes the highlighted word/phrase.
"Begin line"	Moves cursor to beginning of segment.
"Finish line"	Moves cursor to end of segment.
"Next/Previous Segment"	Saves the current segment and then moves to the next/previous segment.
"Save segment"	Saves the current segment.
"Clear segment"	Clears text in the current segment.
"Comma" "Period" "Question mark" "Semicolon" etc.	Inserts punctuation marks ",", ".", "?", ";" in the current cursor location.

Figure 01 shows when the student has opened the PE task, where the source language is French and the target language is English. The student can speak (by clicking the microphone image), type or speak commands for each segment. In a PE task, the MT suggestions that were included when the task was created will appear on the target (right) side. In a translation task, the target side will be empty by default (the MT suggestion can however be made visible below the source segment by clicking a button, if needed.)

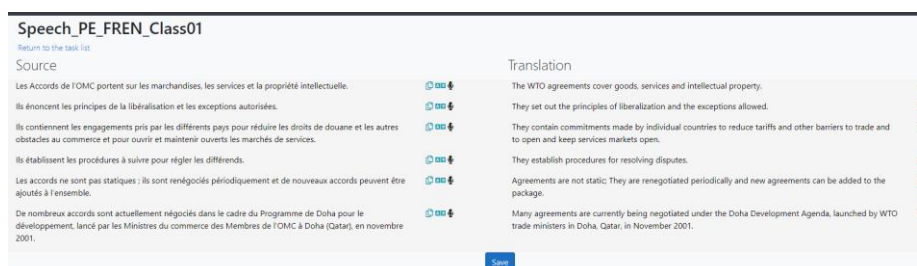


Fig. 1. Student view of a task assigned by the teacher

Figure 02 shows when a student uses spoken post-editing in the second segment. The student has switched the microphone on and had asked via a speech command to select the phrase “exceptions allowed”, which is highlighted. The speech command is printed below, under Speech Commands, with a time log. Once a segment is selected, the student can either delete, replace or undo the selection.

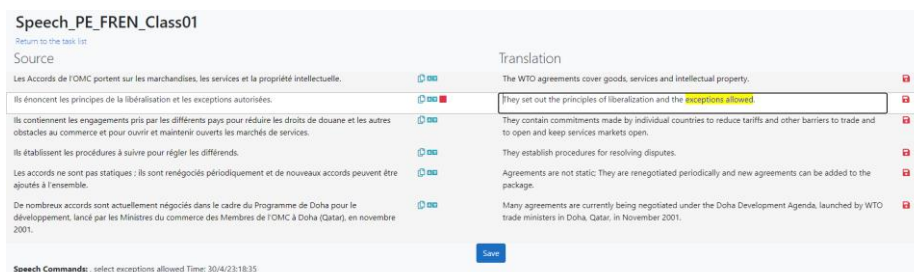


Fig. 2. Using spoken commands to post-edit the segment (e.g. "select exceptions allowed")

As shown in Table 01, some other spoken commands can be used to move in between segments, save segments, clear the entire segment, move to beginning and end of the segment and to speak out punctuation marks. Some other sub commands perform capitalizing the beginning of the sentences/specific words and phrases. All these commands can be developed for other target languages as well upon need.

Once the translations are complete, the student submits the task. The submitted tasks will be sent to the corrector/teacher, along with the statistics of each segment: the number of keystrokes, time taken and average statistics of each task as well. The corrector will then annotate the errors using pre-defined/pre-imported error annotation

schemas. Figure 03 shows the corrector annotating a terminology error in one of the segments.

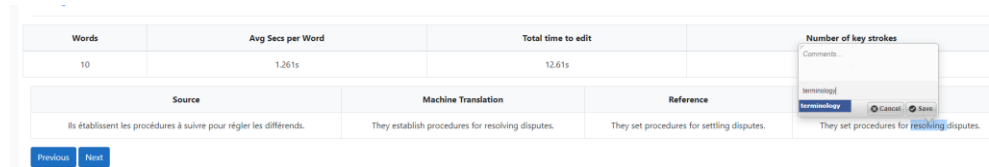


Fig. 3. Corrector annotates a terminology error found in the translation

Once the correction task is complete, detailed information and statistics are provided on the type and description of corrections made. Figure 04 displays one such log, where it shows the reference, source, machine translation target, postedit and the corrections made by the teacher, along with the type of error annotation. These statistics on both translator and corrector side, along with the corpora, allow translation trainers to better understand spoken translation/post-editing behavior.

Errors

Title	Color	Description	#Error
fluency	Orange	Issues related to the form or content of a text, irrespective of whether it is a translation or not.	1
terminology	Red	A term (domain-specific word) is translated with a term other than the one expected for the domain or otherwise specified.	1
fluency spelling	Blue	Issues related to spelling of words.	1
verily completeness	Purple	The text is incomplete.	1

Production

Ref	Source	Target	PEtarget	PEtargetCorrected	SourceTime	TargetTime	Keystrokes	AdditionalTime
The WTO agreements cover goods, services and intellectual property. They set out the principles of liberalization and the permitted exceptions.	Les Accords de l'OMC portent sur les marchandises, les services et la propriété intellectuelle. Ils énoncent les principes de la libéralisation et les exceptions autorisées.	The WTO agreements cover goods, services and intellectual property. They set out the principles of liberalization and the permitted exceptions.	The WTO agreements cover goods, services and intellectual property. They set out the principles of liberalization and the permitted exceptions.	The WTO agreements cover goods, services and intellectual property. They set out the principles of liberalization and the permitted exceptions.	0:00	0:00	0	0:00
They include individual countries' commitments to lower customs tariffs and other trade barriers, and to open and keep open services markets.	Ils contiennent les engagements pris par les différents pays pour réduire les droits de douane et les autres obstacles au commerce et pour ouvrir et maintenir ouverts les marchés de services.	They contain commitments made by individual countries to reduce tariffs and other barriers to trade and to open and keep services markets open.	They include individual countries' commitments to lower tariffs and other trade barriers and to open and keep services markets open.	They include individual countries' commitments to lower tariffs and other trade barriers and to open and keep services markets open.	2:14.52	0:00	10	0:00
They set procedures for settling disputes.	Ils établissent les procédures à suivre pour régler les différends.	They establish procedures for resolving disputes.	They set procedures for resolving disputes.	They set procedures for resolving disputes.	12:51	0:00	4	0:00
These agreements are not static; they are renegotiated from time to time and new agreements can be added to the package.	Les accords ne sont pas statiques. Ils sont renégociés périodiquement et de nouveaux accords peuvent être ajoutés à l'ensemble.	Agreements are not static; they are renegotiated periodically and new agreements can be added to the package.	These agreements are not static; they are renegotiated on to time and new agreements can be added to the package.	These agreements are not static; they are renegotiated on to time and new agreements can be added to the package.	25:11	0:00	11	0:00
Many agreements are currently being negotiated under the Doha Development Agenda, launched by WTO trade ministers in Doha, Qatar, in November 2001.	De nombreux accords sont actuellement négociés dans le cadre du Programme de Doha pour le développement lancé par les ministres du commerce des membres de l'OMC à Doha (Qatar), en novembre 2001.	Many agreements are currently being negotiated under the Doha Development Agenda, launched by WTO trade ministers in Doha, Qatar, in November 2001.	Many agreements are now being negotiated under the Doha Development Agenda, launched by WTO trade ministers in Doha, Qatar, in November 2001.	Many agreements are now being negotiated under the Doha Development Agenda, launched by WTO trade ministers in Doha, Qatar, in November 2001.	23:51	0:00	5	0:00

Fig. 4. Detailed information and statistics on post-editing tasks and error annotation

4 Conclusion

In this paper, we present the different functionalities of COPECO, a speech input enabled online teaching and learning platform developed to collect learner translation/post-editing corpora as well as to learn speech based translation and post-editing behavior and statistics. Future work includes adding new custom voice commands when needed for multiple languages and improving the usability of speech modalities, based on user feedback. While currently we use Google Web Speech API, we can also integrate other ASR services in the future. This would make this platform also able to be used to compare post-editing behaviors (productivity, fatigue and engagement) based on different speech recognizers and different translator profiles.

References

1. Aziz, W., Castilho, S., & Specia, L. (2012, May). PET: a Tool for Post-editing and Assessing Machine Translation. In LREC (pp. 3982-3987).
2. Bolaños-García-Escribano, A., Díaz-Cintas, J., & Massidda, S. (2021). Latest advancements in audiovisual translation education. *The interpreter and translator trainer*, 15(1), 1-12.
3. Bowker L. (2002). *Computer-Aided Translation Technology: A Practical Introduction*. Ottawa, Canada: University of Ottawa Press.
4. Brockmann, J., Wiesinger, C., & Ciobanu, D. (2022, June). Error Annotation in Post-Editing Machine Translation: Investigating the Impact of Text-to-Speech Technology. In *Proceedings of the 23rd Annual Conference of the European Association for Machine Translation* (pp. 249-257).
5. Chrome Web Speech API. <https://developer.chrome.com/blog/voice-driven-web-apps-introduction-to-the-web-speech-api/> last accessed 2023/04/25
6. Ciobanu, D. (2014). Of dragons and speech recognition wizards and apprentices. *Tradumatica* 2014, 524–538. doi: 10.5565/rev/tradumatica.71
7. Ciobanu, D. (2016). Automatic Speech Recognition in the professional translation process. *Translation Spaces. A multidisciplinary, multimedia, and multilingual journal of translation*, 5(1), 124-144.
8. Ciobanu, D., and Secară, A. (2020). “Speech recognition and synthesis technologies in the translation workflow” in *The Routledge Handbook of Translation and Technology*. ed. M. O’Hagan (Milton Park: Routledge), 91–106.
9. Dragon Speech Recognition Solutions, <https://www.nuance.com/dragon.html>, last accessed 2023/04/25
10. Dymetman, M., Brousseau, J., Foster, G., Isabelle, P., Normandin, Y., and Plamondon, P. (1994). Towards an automatic dictation system for translators: the TransTalk project. In *Proceedings of the International Conference on Spoken Language Processing (ICSLP)*, pp. 1–4.
11. Garcia-Martinez, M., Singla, K., Tammewar, A., Mesa-Lao, B., Thakur, A., Anusuya, M. A., & Carl, M. (2014). SEECAT: Speech & Eye-tracking Enabled Computer Assisted Translation. In *European Association for Machine Translation: EAMT* (pp. 81-88).
12. Herbig, N., Düwel, T., Pal, S., Meladaki, K., Monshizadeh, M., Krüger, A., & van Genabith, J. (2020, July). MMPE: A multi-modal interface for post-editing machine translation. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics* (pp. 1691-1702).
13. Herbig, N., Pal, S., Düwel, T., Shenoy, R., Krüger, A., & van Genabith, J. (2020, October). Improving the multi-modal post-editing (MMPE) CAT environment based on professional translators’ feedback. In *Proceedings of 1st Workshop on Post-Editing in Modern-Day Translation* (pp. 93-108).
14. Liyanapathirana, J., Bouillon, P., & Mesa-Lao, B. (2019, August). Surveying the potential of using speech technologies for post-editing purposes in the context of international organizations: What do professional translators think?. In *Proceedings of Machine Translation Summit XVII: Translator, Project and User Tracks* (pp. 149-158).
15. Liyanapathirana, J., & Bouillon, P. (2021). Integrating post-editing with Dragon speech recognizer: a use case in an international organization. *Translating and the Computer* 43, 55-67.

16. Liyanapathirana, J. et al. Integrating Speech in Post-Editing (PE)-Comparison of two PE Interfaces. In: *New Trends in Translation and Technology (NeTTT)*. Rhodes Island, Greece. [s.l.] : [s.n.], 2022. p. 120–123.
17. MateCat Homepage, <https://site.matecat.com/>, last accessed 2023/04/25
18. MateCat Guide, <https://guides.matecat.com/translate-1>, last accessed 2023/04/25
19. memoQ Homepage, <https://www.memoq.com/>, last accessed 2023/04/25
20. Mesa-Lao, B. (2014, April). 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* (pp. 99-103).
21. Mossop B. (2006). Has computerization changed translation? *Meta* 51, 787–805. doi: 10.7202/014342ar
22. Mutal, J.D., Bouillon, P., Schumacher, P. and Gerlach, J., 2020. COPECO: a Collaborative Post-Editing Corpus in Pedagogical Context. In North American component of the International Association for Machine Translation. 1st Workshop on Post-Editing in Modern-Day Translation.
23. O'Brien, S. (2011). Towards a dynamic quality evaluation model for translation. In *Journal of Specialized Translation*, 17:1–2.
24. O'Brien S. (2012). Translation as human–computer interaction. *Transl. Spaces* 1, 101–122. doi: 10.1075/ts.1.05obr
25. Reddy, A., and Rose, R. C. (2010). Integration of statistical models for dictation of document translations in a machine-aided human translation task. *IEEE Trans. Audio Speech Lang. Process.* 18, 2015–2027. doi: 10.1109/TASL.2010.2040793
26. Ruiz, N., & Federico, M. (2014). Assessing the impact of speech recognition errors on machine translation quality. In *Proceedings of the 11th Conference of the Association for Machine Translation in the Americas: MT Researchers Track* (pp. 261-274).
27. Teixeira, C. S., Moorkens, J., Turner, D., Vreeke, J., & Way, A. (2019, March). Creating a multimodal translation tool and testing machine translation integration using touch and voice. In *Informatics* (Vol. 6, No. 1, p. 13). MDPI.
28. TradDICT Learn Homepage, <https://www.traddictlearn.online/>, last accessed 2023/04/25
29. Toral, A., Wieling, M. and Way, A., 2018. Post-editing effort of a novel with statistical and neural machine translation. *Frontiers in Digital Humanities*, 5, p.9.
30. Trados Studio Homepage, <https://www.trados.com/products/trados-studio/>, last accessed 2023/04/25
31. Vilar, D., Xu, J., D'Haro, L. F., and Ney, H. (2006). Error analysis of statistical machine translation output. In *Proceedings of the Fifth International Conference on Language Resources and Evaluation (LREC'06)*, pages 697–702.
32. Wang, L., & Sun, S. (2023). Dictating translations with automatic speech recognition: Effects on translators' performance. *Frontiers in Psychology*, 14.
33. Zapata, J. (2014, November). Exploring multimodality for translator-computer interaction. In *Proceedings of the 16th International Conference on Multimodal Interaction* (pp. 339-343).
34. Zapata, J., Castilho, S., & Moorkens, J. (2017). Translation dictation vs. post-editing with cloud-based voice recognition: A pilot experiment. *Proceedings of MT Summit XVI*, 2.