

**INTERNATIONAL CONFERENCE  
Human-informed Translation  
and Interpreting Technology**

**HiT-IT 2023**

**P R O C E E D I N G S**

Edited by Constantin Orăsan, Ruslan Mitkov,  
Gloria Corpas Pastor and Johanna Monti



**International Conference on  
Human-Informed Translation  
and Interpreting Technology**

7-9 July 2023, Naples, Italy  
<http://hit-it-conference.org>

**INTERNATIONAL CONFERENCE**  
Human-informed Translation  
and Interpreting Technology  
(HiT-IT 2023)

**PROCEEDINGS**

7-9 July 2023, Naples, Italy  
<http://hit-it-conference.org>

Series Online ISSN 2683-0078  
e-book site: [www.acl-bg.org](http://www.acl-bg.org)

INCOMA Ltd.  
Shoumen, BULGARIA

# A New CAI Tool for RSI Interpreters' Training: a Pilot Study

Valentina Baselli <sup>[1]</sup>

<sup>1</sup> IULM University, Milano, Italy

valentina.baselli@iulm.it

**Abstract.** Over the past few years, new technologies in the field of Interpreting have greatly reshaped the way interpreters work, leading to a technological turn in Simultaneous Interpreting (Fantinuoli 2018), due to the increasing use of Remote Simultaneous Interpreting (RSI) and Computer Assisted Interpreting Tools (CAI tools). When there is no human boothmate, AI-based CAI tools are becoming “artificial boothmates” (Fantinuoli 2017), which support the interpreter before and while they deliver Simultaneous Interpreting services through automatic terminology lookup, key term identification, automatic speech recognition, real-time speech transcription, and number highlighting.

While a few researchers have investigated the field of Computer Assisted Interpreting, e. g. Fantinuoli (2017; 2018; 2019), Prandi (2018; 2020), Frittella (2022; 2023) and Defrancq (2020), more research into Computer Assisted Interpreting Training is needed, so that new technologies may be integrated into interpreter training and workflow, given their potential to help interpreters face this technological breakthrough.

This pilot study, conducted within the IULM research project “Collaboration for translation and interpreting: tools and teaching applications”, focuses on investigating the training of interpreting students on these new technologies in collaboration with the RSI-platform Converso Education by integrating the RSI-platform with a new CAI tool specifically developed for teaching purposes.

To the best of our knowledge, this RSI-platform with CAI tool specifically developed for interpreting students based on their needs is the first of its kind.

**Keywords:** AI-powered CAI tool, Remote Simultaneous Interpreting (RSI), Computer Assisted Interpreter Training

## 1 Introduction

Over the past few years, new technologies in the field of interpreting have greatly reshaped the way interpreters work, leading to a technological turn in the sphere of Simultaneous Interpreting (Fantinuoli 2018), as the use of Remote Simultaneous Interpreting (RSI) after the Covid-19 pandemic (Baselli 2023) and Computer Assisted Interpreting tools (CAI tools) has greatly increased.

With the recent integration of CAI tools into RSI platforms, such as Kudo’s Assist and SmarTerp, the development of new tools now aims to increase the efficiency of the interpreter’s workflow and provide interpreters with a better user experience (Frittella 2023).

In 2022, we started teaching a new Remote Simultaneous Interpreting class at IULM University using the Converso Education Platform. In substance, fifty first-year and

fifty second-year students of the Master’s Degree Course in Conference Interpreting attended ten Remote Simultaneous Interpreting lessons.

Besides experiencing technical problems due to poor internet connections or inadequate equipment (such as devices and headphones), we noted that the students also ran into some difficulties during the lessons with the remote interpreting itself, especially where numerals and specialized terms were concerned. For this reason, we asked the students to complete a questionnaire on the main difficulties they encountered during Remote Simultaneous Interpreting and on useful resources to overcome those difficulties. The aim was to establish if the use of a CAI tool including an “artificial boothmate” (Fantinuoli 2017), which displays what are known as "SI problem triggers", might be helpful during RSI lessons.

## 2 Survey on Students’ Requirements

The goal of the survey was to explore the current students’ requirements in the field of computer-assisted interpreter training, to find ways to help trainee conference interpreters face the above-mentioned technological turn, and provide them with the proper tools to adequately manage RSI through the development of a new CAI tool based specifically on their requirements.

### 2.1 Sample

The survey was hosted on Google Forms and sent to the participants via email in April 2023. The thirty participants were regular, full-time students enrolled in the first year of the Master’s Degree Course in Conference Interpreting at IULM University. The participants’ A language was Italian, and B language was English.

### 2.2 Questionnaire

The user requirements for our new CAI tool stem from a questionnaire completed by thirty trainee interpreters at the end of their RSI lessons and a focus group consisting of six students conducted a few days before the recordings were made.

In the questionnaire, the students were asked to answer specific questions on their preferences related to a CAI tool developed for teaching purposes.

**Table 1.** Question 1: If it were possible to receive support during Simultaneous Remote Interpreting, would it be useful for you to see the numbers uttered by the speaker?

Response	Percentage of responses
Yes	97%
No	3%

Almost all the students replied that it would be useful to see the numbers uttered by the speakers on the screen and most of them (67%) preferred the numerals and punctuation to be converted into target language format.

**Table 2.** Question 2: If it were possible to receive support during Simultaneous Remote Interpreting, would it be useful for you to see the specialized terms uttered by the speaker?

Response	Percentage of responses
Yes	93%
No	7%

Almost all the interviewees replied that it would be useful to see the specialized terms uttered by the speakers on the screen and 90% of them would prefer to have them displayed both in the source and target languages.

**Table 3.** Question 3: If it were possible to receive support during Simultaneous Remote Interpreting, would it be useful for you to see the named entities (places, persons, etc.) uttered by the speaker?

Response	Percentage of responses
Yes	93%
No	7%

Among the survey respondents, almost all replied that it would be useful to see the named entities uttered by the speakers on the screen.

**Table 4.** Question 4: If it were possible to receive support during Simultaneous Remote Interpreting, would it be useful for you to see the entire transcription of the speech?

Response	Percentage of responses
Yes	50%
No	50%

It is interesting to note that exactly half of the students would find it useful to see the whole transcript while the other half would prefer not to see the whole running transcription of the speech uttered by the speaker, but only single elements.

### 2.3 Results

As most of the respondents who took part in the survey reported that displaying numbers, specialized terms, and named entities on the RSI platform during RSI classes

would be of great benefit, we decided to develop a CAI tool together with Converso, implementing it in the Converso Education Platform.

This tool is a prototype of an ASR-supported CAI tool that transcribes speech delivered in English and Italian, when enabled by the teacher based on the specific needs, and automatically provides the interpreter with numerals and their unit of measurement, and translation options for terminology (drawn from a previously provided glossary). The Converso Education Platform includes a button called “view list”, which activates either automatic transcription or the CAI tool, and a “focus mode” button, which displays the transcript with highlighted specialized terms and numbers. A two-second latency was chosen, in accordance with studies conducted by Fantinuoli and Montecchio (Fantinuoli and Montecchio 2022), considering the average ear-voice span of interpreters.

Furthermore, according to the students’ preferences and Frittella’s recommendation (Frittella 2023), it was decided to let the CAI tool display numerals in their final version, without the partial rendition proposed by the ASR, as well as the numeral together with the following element in the sentence, which is usually either the referent or the unit of measurement. The above-mentioned items remain on screen for as long as there is enough room on the screen (EABM 2021). In its current state, however, our CAI tool prototype displays the numbers in the source language without showing the following element in the sentence. Adjustments will be made in the coming months before the study is conducted in the Autumn.

### **3 Pilot Study**

#### **3.1 Development of the CAI Tool Integrated in the RSI Platform Converso Education**

According to the principles defined by Fantinuoli (Fantinuoli 2017) for ASR-based CAI tools, the final version of the CAI tool integrated in the Converso Education Platform will:

- be speaker-independent
- be able to manage continuous speech
- support large-vocabulary recognition
- support vocabulary customization for the recognition of specialized terms
- have high performance accuracy, i.e. a low word/error rate (WER)
- have high precision, i.e. fraction of relevant instances among the retrieved instances
- have high recall, i.e. the fraction of relevant instances that have been retrieved over the total amount of relevant instances present in the speech (with precision having priority over recall, in order to avoid producing results that are not useful and may distract the interpreter)
- have a distraction-free graphical user interface to present the results.

### 3.2 Evaluation of the CAI Tool

According to Frittella’s methodology, our ASR-supported CAI tool prototype was evaluated via tool performance, users’ performance, and users’ perception (Frittella 2023: 55).

Tool performance was assessed by adopting the same principle first used by Fantinuoli (Fantinuoli 2017), through accuracy, precision, and recall. Accuracy (i.e. word-error rate) is the percentage of wrongly displayed items (numerals, terminology, and named entities) out of all items that should have been displayed. Precision refers to the number of correct positive results divided by the number of all positive results, and recall indicates the number of correct positive results divided by the number of positive results that should have been returned.

User performance was assessed both through the rendition of individual items (interpreted specialized terms and numerals) and by considering the meaning of the interpreted part of the speech.

Users’ perception was evaluated through a post-task questionnaire.

### 3.3 Preliminary Test

In order to assess tool performance, a preliminary test was carried out with a pre-recorded speech and no interpreting. The aim was to evaluate the ASR precision regarding numbers and terminology. Named entities are not recognized by the CAI tool prototype at this stage.

The development of the tool has taken the principles defined by Fantinuoli (2017) for ASR-based CAI tools into consideration. Specifically, in order to be used with a CAI tool, an ASR system needs to be speaker-independent, be able to manage continuous speech, support large-vocabulary recognition and vocabulary customization for the recognition of specialized terms, and have high performance accuracy.

The table below shows the results of the preliminary test conducted on the specialized terms provided through a glossary and identified by the CAI tool. The speech (described below) was the same interpreted by the subjects during the recordings.

**Table 5.** Results of the preliminary test conducted on specialized terms.

Total specialized terms	Errors	Omissions
44	-	4

The four omitted specialized terms were two noun phrases consisting of five constituents (solar-thermal heating water system, natural-gas-based electricity generation), and two noun phrases containing acronyms (REs installations, Refuse-derived fuel RDF).

As far as numerals are concerned, the following table shows the results of the preliminary test conducted on the numbers identified by the CAI tool.

**Table 6.** Results of the preliminary test conducted on numbers.

Total numbers	Errors	Omissions
21	1	-

The numeral that was incorrectly displayed was 405, which was split into two numbers (400 and 5). The following table shows an overview of accuracy, precision, and recall considering all 65 stimuli (numbers + terms). Precision and recall values are expressed from 0 to 1, with 1 being the maximum value, whereas accuracy is expressed as a percentage of error (the lower the percentage, the more accurate the result).

**Table 7.** Tool performance assessment

WER (accuracy)	Precision	Recall
7.7%	0.98	0.92

### 3.4 Pilot Study

The preliminary test and the pilot study were carried out at the Converso Hub in Milano at the beginning of June 2023. The Converso hub is Italy's first Remote Simultaneous Interpreting hub consisting of fourteen booths, a conference room and three control rooms. Every booth is fitted with professional equipment (ISO 20109:2016), ISO compliant interpreting console, a 27" full-HD display (single monitor in single-desk booths and double monitor in double-desk booths), professional gooseneck microphone or professional headset microphone, and professional headphone.

The six participants were regular, full-time students enrolled in the first year of the Master's Degree Course in Conference Interpreting at IULM University who were attending the RSI classes. The participants' A language was Italian, which is the target language of the study, and their B-language was English, the source language of the speeches.

As usability is determined according to the relevance of a product for a particular user and aim, the participants are representative of the target users (Frittella 2023: 20), that is to say Master's Degree students in Conference Interpreting.

Each test subject was asked to interpret two speeches of a similar length and lexical density on renewable energy transition. In one case, a Microsoft Word table glossary with the relevant terminology was provided. No glossary was provided in advance to help the interpreters with the second speech, but specialized terms and numbers were displayed by the CAI tool. The second speech was reinterpreted by the subjects while the entire ASR transcript of the speech was displayed. The performance of the students with ASR transcript will be compared with that obtained without CAI tool in a future study. A few days before the experiment the participants were given access to the Converso platform with transcription and CAI tool, in order to avoid the "novelty effect". However, according to the post-task questionnaire results, more in-depth training would have been useful.



The two speeches with similar difficulty levels contained the same number of specialized terms and numerals. They have been prepared by the author and pre-recorded by a native American English teacher. Both are interpreter trainers. The average speed of the speeches was 100 words per minute in accordance with the indication given by Korpala and Stachowiak-Szymczak (2020) on the ideal speech rate for interpreters. The first speech (1011 words) was ten minutes and five seconds in duration, with 65 stimuli (21 numbers + 44 terms: one of which was a unigram, 19 bigrams, 20 trigrams, 2 4-grams and 2 5-grams) whereas the second speech (1000 words) was exactly ten minutes long with 65 stimuli (21 numbers + 44 terms: 9 of which were unigrams, 23 bigrams, 10 trigrams, 2 4-grams and no 5-grams, since the CAI tool did not recognize the 5-grams).

After the test, participants were asked to complete a post-task questionnaire on their perception and assessment of the tool and their preference for a display format (terms on the left and numbers on the right or vice versa, new items in a bold font, a larger font size, etc.) in addition to further comments and open questions on the use of the CAI tool.

Subsequently, the subjects' deliveries were checked for the percentage of terms and numbers translated in the first and in the second speech, which indicates whether the use of the CAI tool would help improve terminological and number precision in RSI classes. The following tables show the number of correctly translated stimuli with the support of the CAI tool and with the Microsoft Word glossary.

**Table 8.** Number of correctly translated stimuli with CAI tool

	Student 1	Student 2	Student 3	Student 4	Student 5	Student 6
Numbers	61.9%	81%	95.2%	52.4%	81%	71.4%
Specialized terms	68.2%	54.5%	86.4%	75%	88.6%	63.6%
Total	66.15%	63%	89.2%	67.7%	86.15%	66.15%

When supported by the CAI tool, the six subjects correctly translated 73% of the stimuli on average. Interestingly, the numeral wrongly displayed by the CAI tool (405) was correctly interpreted by two trainees.

**Table 9.** Number of correctly translated stimuli without CAI tool

	Student 1	Student 2	Student 3	Student 4	Student 5	Student 6
Numbers	38%	52.4%	66.7%	28.6%	33.3%	33.3%
Specialized terms	27.3%	43.2%	61.4%	50%	45.5%	29.5%
Total	30.8%	46.15%	63.1%	43.1%	41.5%	30.8%

Conversely, when interpreting with a Microsoft Word glossary, the six subjects correctly translated 42.6% of the stimuli on average. What clearly emerges from the data analysis is that all subjects show a remarkably better and more precise performance in

RSI with CAI tool. However, some interpreted sentences in two trainees' deliveries did not make sense. Although the numbers and specialized terms were correct when supported by the CAI tool, the segments following the numbers were sometimes mistranslated. A qualitative analysis relating to the meaning of the interpreted speech segments will be included in the future study.

Comparing the two tables, it can be observed that the number of correctly interpreted stimuli (both numerals and specialized terms) is higher in the Remote Simultaneous Interpreting with CAI tool support than in the RSI performed with Microsoft Word glossary for all six subjects (student 1: 66.15% vs 30.8%; student 2: 63% vs 46.15%; student 3: 89.2% vs 63.1%; student 4: 67.7% vs 43.1%; student 5: 86.15% vs 41.5%; and student 6: 66.15% vs 30.8%). The data from the future study on a larger sample with the adjusted CAI tool will provide a broader view of the phenomenon and produce more findings regarding the students' deliveries with or without CAI tool.

### 3.5 Results from the Post-Task Questionnaire

The trainees were asked to complete a post-task questionnaire after the test to evaluate user perception and satisfaction. According to the results, the 6 subjects were overall satisfied with the use of the CAI tool and emphasized that it is easy to use and intuitive, but some adjustments need to be carried out to make the CAI tool even more effective, as it is a prototype. The average scores obtained in the various categories analyzed (perceived ease of use, effectiveness, ease of learning, timeliness, dependability) range from 6.8 to 8.8, with a prevalence of an average score of 8 out of 10.

**Table 10.** Users' perception

Question	Average of the 6 scores (from 1 to 10)
Your satisfaction with the CAI tool's support during the test	8
The CAI tool was easy to use	8.8
The CAI tool helped me improve the accuracy of my delivery	8
No training is required to use the CAI tool effectively	6.8
The input provided by the CAI was timely	7.8
I felt that I could rely on the CAI tool's support	8

In terms of design-related recommendations, subjects have different opinions on the choice of graphic representation of the elements displayed by the CAI tool. Currently, specialized terms, numbers, and transcriptions are shown in a box on the right-hand side of the screen while new items appear in the same font and remain on the screen until there is no more room.

**Table 11.** Design-related recommendations given by the subjects.

Specialized terms	on the left 2	on the right 3	In the center 1
Numbers	on the left 1	on the right 3	In the center 2
New items	bold font 5	larger font size 0	No font change 1

Moreover, it is interesting to note that 5 out of 6 stated that they perceived the CAI tool as being more reliable than a human boothmate in terms of precision and speed in suggesting specialized terms and numbers. On the one hand, it emerged that for a trainee it is reassuring to know that the interpreter has not to search for terminology through a glossary, even if a human boothmate psychologically helps make the interpreter feel less alone and can better understand the interpreters' needs. On the other hand, a trainee found the transcript and particularly the stimuli moving on the screen very distracting. It would be interesting to see if the same difficulties would emerge after more training sessions with the CAI tool.

#### **4 Conclusions and Future Work**

This test is a pilot study for a broader research project of the “International Center for Research on Collaborative Translation - IULM” that involves the development of a CAI tool integrated into the Converso Education RSI Platform and the usability test. The findings of this pilot study will help us redesign the broader study which is scheduled for Autumn 2023 on a larger sample (between thirty and fifty students enrolled in the second year of the Master's Degree Course in Conference Interpreting at IULM University). Furthermore, the future study will investigate ASR output in the form of both “short prompts” and full ASR transcript, since the author has also decided to record the students' performance with the entire speech transcript. This comparison might lead to more findings which may complement the existing insights from Fantinuoli, Frittella, and Prandi. The results of this pilot study are not intended to be exhaustive but form the basis for in-depth research work on Computer Assisted Interpreting Training.

#### **Acknowledgment**

The author wishes to thank Converso for developing the CAI tool.

## References

1. Baselli, V.: La traduzione collaborativa in interpretazione: il connubio tra interprete e tecnologia. In: Laurenti, F. (ed.), *La traduzione collaborativa tra didattica e mercato globale delle lingue*, Aracne, Roma (2022).
2. Baselli, V.: L'Interpretazione Simultanea da Remoto: stato attuale, sfide e prospettive. In: Fuehrlinger, S., Proietti, P. (eds.), *Transiti linguistici - Traduzioni, interpretazioni, mediazioni linguistiche*, Aracne, Roma (2023).
3. Converso Homepage, <https://converso.cloud/en/> last accessed 2023/4/20.
4. Defrancq, B., Fantinuoli, C.: Automatic Speech Recognition in the booth: Assessment of system performance, interpreters' performances and interactions in the context of numbers. *Target*, DOI (2020).
5. Ergonomics for the Artificial Booth Mate (EABM) <https://www.eabm.ugent.be/survey/> (2021).
6. Fantinuoli, C.: Speech recognition in the interpreter workstation. In: Esteves-Ferreira, J., Macan, J., Mitkov, R., Stefanov, O. (eds.), *Proceedings of the Translating and the Computer 39*, 25–34. AsLing, London (2017).
7. Fantinuoli, C.: Interpreting and technology: The upcoming technological turn. *Interpreting and technology*, Language Science Press, Berlin (2018a).
8. Fantinuoli, C.: Computer-assisted interpreting: challenges and future perspectives. In: Durán, I., Corpas, G. (eds.), *Trends in e-tools and resources for translators and interpreters*. 153-174, Leiden, Brill (2018b).
9. Fantinuoli, C.: The technological turn in interpreting: the challenges that lie ahead. In: *Proceedings of the conference Übersetzen und Dolmetschen 4.0. - Neue Wege im digitalen Zeitalter*, 334-354 (2019).
10. Fantinuoli, C., Prandi, B.: Teaching information and communication technologies: a proposal for the interpreting classroom. *Journal of Translation and Technical Communication Research*, 11(2), 162-182. (2018).
11. Fantinuoli, C., Montecchio, M.: Defining maximum acceptable latency of AI-enhanced CAI tools. In: *Proceedings of LingTech21* (2022).
12. Frittella, F.: CAI Tool-Supported SI of Numbers: A Theoretical and Methodological Contribution, *International Journal of Interpreter Education*: Vol. 14: Iss. 1, Article 5. (2022).
13. Frittella, F.: Usability research for interpreter-centred technology: The case study of SmarT-erp. (*Translation and Multilingual Natural Language Processing 21*). Language Science Press, Berlin (2023).
14. Korpál, P., Stachowiak-Szymczak, K.: Combined problem triggers in simultaneous interpreting: exploring the effect of delivery rate on processing and rendering numbers. *Perspectives Perspectives: Studies in Translation Theory and Practice*, 28 (1): 126–143, <https://doi.org/10.1080/0907676X.2019.1628285> (2020).
15. Prandi, B.: Designing a Multimethod Study on the Use of CAI Tools during Simultaneous Interpreting. In: *Proceedings of the 39th Conference Translating and the Computer*. Translating and the Computer, London (2017).
16. Prandi, B.: An exploratory study on CAI tools in Simultaneous Interpreting: theoretical framework and stimulus validation. In: Fantinuoli, C. (ed.) *Interpreting and technology*, pp. 28–59, Language Science Press, Berlin (2018).
17. Prandi, B.: CAI tools in interpreter training: where are we now and where do we go from here?, in *TRAlinea* (2020).