Sino-US English Teaching, November 2022, Vol. 19, No. 11, 405-412
doi:10.17265/1539-8072/2022.11.004

A Review on Research on Technology-Assisted Simultaneous Interpreting

XU Qianqian
Shanghai Institute of Technology; Shanghai International Studies University, Shanghai, China

Simultaneous interpreting (SI) has undergone a long historical period. With a variety of technologies employed in SI, the researches on it have taken on some new features and brought about new inspirations to this industry. This paper aims at first introducing various technologies or technological tools used in interpreting, and then reviewing on the history of the development of technology-assisted interpreting, the current researches, and their limitations, and finally drawing some inspirations from the past researches of technology-assisted SI, to offer some proposals on the future research trends on technology-assisted SI, ranging from carrying out more empirical studies and surveys to collect interpreters' attitudes towards technology used in SI, to conducting psychological experiments to analyze and explain interpreters' emotional response to technology-assisted SI, to designing specific measurements and indicators to reveal interpreters' discomfort experienced during technology-assisted SI.

Keywords: simultaneous interpreting, technology-assisted, application, training

Introduction

Although technologies have been influencing interpreting in different settings and of different modes, with the emergence of professional interpreting in the 20th century, they did exert different influences on different modes of interpreting in different settings, for the development of different technologies naturally determined the arrival of different interpreting modes. Progress in electro-technical communication equipment facilitated and contributed to the emergence of simultaneous interpreting (SI) in conference settings and resulted in conference interpreting as a profession. However, application of technology in community interpreting or public service interpreting was as late as the last decades, especially in legal interpreting and healthcare interpreting. In addition, recent advances in information and communication technology (ICT) are bringing about fundamental changes to interpreting practices by enabling remote interpreting without interpreter's physical presence onsite (Svoboda, 2014; Władyka, 2014, as cited in Pöchhacker, 2015, p. 410). Finally, the growing use of videoconference technology and remote interpreting has led to considerable researches on automatic speech-to-speech translation, or machine interpreting (MI) (Pöchhacker, 2015, p. 410).

Despite of the progress of MI and these advantages of technologies in interpreting training, there are also some limitations, among which is the limited scope of MI to specific domains and linguistic contexts and to a narrow range of highly standardised natural speech inputs (Jekat, 2000, as cited in Pöchhacker, 2015, p. 242). On

Fund: A Corpus-Based Study on Interpretation Strategies, a project for liberal arts funded by Shanghai Institute of Technology, (Project No.: 10210Q180045036).
XU Qianqian, lecturer, School of Foreign Languages, Shanghai Institute of Technology; doctoral student, Graduate Institute of Interpretation and Translation, Shanghai International Studies University, Shanghai, China.

the other hand, there are also negative comments on technology's impact on SI. Interpreters have seen them as posing a further threat to their visibility, even though the quality of performance achieved in remote mode may not be significantly lower than when working on site (Roziner & Shlesinger, 2010, as cited in Pöchhacker, 2015, p. 347).

Many researches have been done on technology-assisted SI with a variety of research topics ranging from interpreting quality to technological benefits, etc. Therefore, this paper attempts to summarize the limitations of the current researches on technology-assisted SI by reviewing on the history of the development of technology-assisted interpreting and the existing researches, in order to draw some inspirations from the current researches of technology-assisted SI, and finally offer some thoughts on the future research trends of technology-assisted SI.

Literature Review

Review on the Historical Development of Technology-Assisted Interpreting

Developing stages. Sandrelli defines Computer Assisted Interpreter Training (CAIT) as the practice of application of computer technology to enhance interpreter training (Sandrelli, 2007, as cited in Pöchhacker, 2015, p. 75). The history of CAIT can be divided into three stages: early period, middle period, and current development period. The early period is characterized by the application of digital speech banks, and interpreting courseware, computer-based teaching packages, to interpreting training. Its advantage lies in the convenience for teachers without specific software development skills to create their own CAIT materials by integrating various forms of materials, ranging from audio, video to textual resources into interpreting exercises, modules and courses (Sandrelli, 2007; Sandrelli & Manuel Jerez, 2007, as cited in Pöchhacker, 2015, p. 75). Middle period features the expansion of the range of possible CAIT applications with online collaborative activities encouraging user interaction and collaborative learning (Sandrelli, 2007, as cited in Pöchhacker, 2015, p. 76). In current development period, interpreting training beyond the classroom is not only made possible, but also becomes more immersive and realistic with professional settings (Sandrelli, 2007, as cited in Pöchhacker, 2015, p. 75). Zhang et al. also reviewed the history of machine translation, starting with rule-based machine translation, then statistics-based translation, and eventually to neural-system-like machine translation—that is, neural translation based on Deep Learning (Zhang, Yang, Liu, & Li, 2018, pp. 93-94).

Review on the application of technologies to interpreting. Many advanced technologies have been used in interpreting, and different technologies featuring different functions will facilitate interpreting in different ways. The frequently used technologies or technological tools in interpreting include videoconferencing technology, information communication and technology (ICT), web-based video platforms (e.g., YouTube), Speech Repository by the European Commission, speech banks, digital recording technologies which can capture authentic speech material and also record student output in the classroom in dual-track mode (Kalina & Ziegler, 2015, as cited in Pöchhacker, 2015, p. 411), eye tracking, automatic speech recognition (ASR) systems, and so on. A typical example of the digital recording technology is the digital pen or Smartpen, a writing instrument for audio recording as well as the capture and replay of handwritten notes. It can not only be used in self-study in dialogue and consecutive interpreting, but also be employed in simultaneous consecutive interpreting, allowing for note-taking and audio recording at the same time (Kalina & Ziegler, 2015, as cited in Pöchhacker, 2015, p. 412). Additionally, virtual reality systems have also employed highly complex technology for interpreting training, which enables students to interact as avatars in virtual environments simulating specific settings of

professional practice (Braun et al., 2013, as cited in Pöchhacker, 2015, p. 412). Moreover, a variety of applications including software and electronic tools have been demonstrating different facilitating effect on interpreting, some for the documentation and management of terminology in the preparation phase, while others for providing authoritative reference of terminology translation from professional conference interpreters. Software with such function served as database tools or workbench systems, such as Interplex and InterpretBank (Fantinuoli, 2006, as cited in Pöchhacker, 2015, p. 412), which may also incorporate term extraction capabilities (Fantinuoli, 2006, as cited in Pöchhacker, 2015, p. 412). In addition to software, appropriate hardware and mobile devices like notebook and tablet computers have become an essential part of interpreters' working environment (Kalina & Ziegler, 2015, as cited in Pöchhacker, 2015, p. 412).

A very successful application of technologies in interpreting is MI. As for the working mechanism of MI, the typical architectures of MI systems combine ASR and text-to-speech modules with machine translation (MT), which could be considered to be inadequate, because it fails to account for the paraverbal and nonverbal features of spoken language, such as prosody, emphasis, and gestures. However, improvements have been made on the MI system by including additional components, such as processing of context, noise reduction, and dialogue semantics (Jekat, 2000, as cited in Pöchhacker, 2015, p. 239). Working mechanism of MI can be further divided into different categories, including text-to-text (MT) system based on the analysis of text in the corpus, text-to-speech system, and framework integrating ASR and text-to-text system (TTS) (Jekat, 2000, as cited in Pöchhacker, 2015, pp. 240-241). One of the first European MI projects, an important milestone, was Verbmobil (1993-2000), a speech-to-speech translation system for face-to-face dialogue (Wahlster, 2000, as cited in Pöchhacker, 2015, p. 241) in which large databases of spontaneous and semi-structured dialogues were compiled (Jekat & Hahn, 2000, as cited in Pöchhacker, 2015, p. 241). For MI systems processing several languages, the current challenges are automatic speech identification, multilingual acoustic models, and multilingual language models (Pöchhacker, 2015, p. 240).

Other technologies applied in interpreting include eye tracking and ASR. Eye tracking refers to the process of measuring the position and the movement of the human eye with devices known as eye trackers (Pöchhacker, 2015, p. 157), which can be applied to research on various different topics in interpreting, such as the exploration of the multimodal nature of the interpreting process (Pöchhacker, 2015, p. 157), measurement of the cognitive load in SI by using pupil dilation (Seeber, 2013, as cited in Pöchhacker, 2015, p. 157), the use of visual-verbal and visual-spatial information during SI (Seeber, 2012, as cited in Pöchhacker, 2015, p. 157), and so on. ASR is divided into four steps: feature extraction which extracts relevant features from the signal and delivers a set of feature vectors; unit matching which matches feature vectors to sound forms; lexical decoding which compares a sequence of sounds or syllables with entries in the lexicon; syntax analysis which checks whether the recognised sequence of words maps to grammatical constraints of the language model or data in the language model (Bourlard et al., 2011; Haiber, 2004, as cited in Pöchhacker, 2015, p. 240).

Review on the Research Topics

The current research topics on technology-assisted SI can be divided into the following categories. Firstly, researches on the technologies applied in SI from three different dimensions, namely, technology for rendering interpreting services; technology for interpreter training; and technology to aid an interpreter's performance (Kalina & Ziegler, 2015, as cited in Pöchhacker, 2015, pp. 410-412). Secondly, researches on interpreters' specific needs and practices of the documentation and management of terminology (Stoll, 2009; Jiang, 2013, as

cited in Pöchhacker, 2015, p. 412). Thirdly, researches are also done on remote interpreting, which is one of the most important research areas assisted by technologies. The research topics of remote interpreting can be further divided into feasibility studies, interpreting quality or interpreter's performance, interpreter's emotional response to remote interpreting (RI), interpreters' sense of discomfort with RI (Braun, 2013, as cited in Pöchhacker, 2015, pp. 346-347). The feasibility studies, always done in the simultaneous mode of RI, use various technical conditions to reveal physiological and psychological challenges (Braun, 2013, as cited in Pöchhacker, 2015, p. 346). Interpreters' performance in RI is where researchers may disagree with each other, with some studies revealing that interpreters' performance in RI declined faster than their on-site performance, and others revealing no significant differences in RI and on-site performance (Braun, 2013, as cited in Pöchhacker, 2015, p. 347). Consensus has not been achieved on the quality of RI. Studies on RI in legal settings also revealed this discrepancy on the quality; moreover, it is impossible to say without reservation that training, familiarization, and the use of better equipment led to a clear improvement in performance (Braun & Taylor, 2015, as cited in Pöchhacker, 2015, p. 347). Interpreters' emotional response to RI is studied through measuring stress indicators and aspects of working environment (Braun, 2013, as cited in Pöchhacker, 2015, p. 347). Interpreters show their preference among different modalities of interpreting. Azarmina and Wallace note that interpreters generally preferred on-site interpreting to RI, and video to telephone (Azarmina & Wallace, 2005, as cited in Pöchhacker, 2015, p. 348). Although a sense of discomfort with RI on the part of the interpreters is recognized in the studies, it is difficult to account for by objective measures (Roziner & Shlesinge, 2010, as cited in Pöchhacker, 2015, p. 347). In addition, interpreters may even show their reluctance to use technology. Despite various efforts at developing technology for note-taking in an electronic format, or for storing source text information and related details in some other way, interpreters continue to take notes with pen and notepad (Kalina & Ziegler, 2015, as cited in Pöchhacker, 2015, p. 411).

In addition, researchers from China have also conducted researches on a variety of topics in this field. Wang and Wang has conducted a research on major platforms such as KUDO, Interprefy, VOICEBOXER, Interactio, SPEAKUS, VERSPEAK, which summarized the history of the development of technologies for remote interpreting and compared the functions, user-friendliness, features, and limitations of remote simultaneous interpreting (RSI) (X. M. Wang & B. H. Wang, 2021, p. 105). Li and Wang (2018) have conducted an empirical research on the construction of the teaching model by using ASR APP iFLYTEK in SI training course, which consists of five stages: speaker, speech recognition APP, interpreter, production (sight interpreting), and audience (p. 14) and has the following benefits: promoting the completeness of translation and accuracy of term translation, making SI task easier, and increasing students' interest in learning interpreting (p. 16). However, the limitations of the speech recognition APP of iFLYTEK lie in distracting students' attention, low accuracy in recognizing words with similar pronunciation (Li & Wang, 2018, p. 17). Lu has explored the feasibility and effectiveness of online interpreting in a SI course via video conferencing at GSTI, BFSU, with surveys done to collect students' feedback on the effectiveness, advantages and disadvantages of the course. However, Lu pointed out some deficiencies of the new teaching paradigm, as well as the challenges to the application of distance mode (Lu, 2020, p. 191). Lu has reviewed the researches on SI during 40 years from 1980 to 2020 and emphasized that in its comprehensive development stage, researches on SI are characterized by their technological perspectives with media SI, multimodal corpus, distant learning, speech recognition, machine assisted SI teaching, and practice drawing growing attention (Lu, 2022, p. 35). Sun, Li, and Lu (2021) has conducted an empirical research on AI-assisted SI, which transcribed the original speech text on the computer screen and also showed a machine-

translated version, to explore its benefits and potential negative impacts on interpreters' performance (p. 12). Zhang et al. (2018) have conducted a research on human-computer collaboration through developments in artificial intelligence, which summarized the features of machine translation and put forward the future improvements to be made, by comparing the results between manually translated and machine-translated samples (p. 93).

Reflections on Current Researches and Proposals for Future Research Topics

Reflections on Current Researches

Based on the above analysis of the current researches on technology-assisted SI, the author expects to make some reflections on current researches from both their focus on technological perspectives and the imbalanced research topics related to technology-assisted SI. Once, interest in the strictly technological aspects remained rather limited in the field of interpreting (Kalina & Ziegler, 2015, as cited in Pöchhacker, 2015, p. 411), but now there seems to be many researches with too much focus on technological side instead of on interpreting itself. Taking all the positive and negative implications of technology in SI teaching and practice concluded from the current researches, this paper aims at proposing some topics for further research in the domain of technology-assisted SI.

First, an imbalance of research topics can be deducted from the analysis of the current researches, the majority of which focus on such topics as technological benefits (Sun et al., 2021, p. 12), the working mechanism of technology-assisted SI, etc., with a few researches delving into the negative impacts of technology on SI. However, the technological benefits discussed are limited to a few aspects such as helping interpreters with terminologies in interpreting tasks (Zhang et al., 2018, p. 92), assisting interpreters in pre-event preparation (Sun et al., 2021, p. 12), and are more for translators than for interpreters (Zhang et al., 2018, p. 92). Although some researches discussed the negative impacts of technology on SI, they are limited to causing distraction to interpreters, adding extra burden, and making the logic connection more difficult (Sun et al., 2021, p. 12).

Second, there are some limitations of researches on SI from technological perspectives. Despite the facilitation of technology in SI research, researches on SI from technological perspectives and by technological methodologies may have some limitations. According to Pöchhacker, eye tracking, as a research method in interpreting studies, is objective and has limited invasiveness to the ongoing SI task, which features high temporal resolution to time-lock ocular phenomena and cognitive processes down to the level of phonemes (Pöchhacker, 2015, p. 157). However, there are also some limitations of this technological method, among which are covert attention (Wright & Ward, 2008, as cited in Pöchhacker, 2015, p. 157) detracting from the eye-mind hypothesis, and many measures with noteworthy exceptions, such as pupil dilation, which should be dealt with by a visual input component (Wright & Ward, 2008, as cited in Pöchhacker, 2015, p. 157).

Proposals for Future Research Topics

With the widespread and increasing employment of technologies in SI, potential research topics do vary, and some researchers continue to show their research interest in technological endowment in this field. Lu insisted that not many researches on technology used in SI were conducted, and thus proposed that a SI research laboratory should be built to carry out research on the cognitive process of SI with the help of modern technical means, such as eye tracker, near infrared spectrum scanner, Electroencephalogram (EEG) scanner, magnetic resonance imaging, and a machine SI system to carry out comparative and collaborative research on human-

machine SI (Lu, 2022, pp. 36-37). Moreover, rather marginal interest has been devoted to technological improvements in portable equipment for SI, possibly because of conference interpreters' reluctance to use it especially when working in booths where mobility is not a requirement (Pöchhacker, 2015, p. 411).

Although technologies have greatly reshaped SI and lots of researches have been done on technology-assisted SI, many of the most studied topics include interpreters' performance, the working mechanism of technology-assisted SI, etc., without taking into account interpreters' subjective attitude and motivation in using technology in SI. Therefore, more researches can be done on such topics as interpreters' emotional response to technology-assisted SI, interpreters' preference among different modalities, on-site SI or technology-assisted RI, interpreters' discomfort experienced during technology-assisted SI, interpreters' reluctance to use technology. Currently, there are a few researches that have been done on these topics. Interpreters' emotional response to RI is studied through measuring stress indicators and aspects of working environment (Braun, 2013, as cited in Pöchhacker, 2015, p. 347). Interpreters show their preference among different modalities of interpreting. Azarmina and Wallace note that interpreters generally preferred on-site interpreting to RI, and video to telephone (Azarmina & Wallace, 2005, as cited in Pöchhacker, 2015, p. 348). Although a sense of discomfort with RI on the part of the interpreters is recognized in the studies, it is difficult to account for by objective measures (Roziner & Shlesinger, 2010, as cited in Pöchhacker, 2015, p. 347). Despite various efforts made at developing technology for note-taking in an electronic format, interpreters continue to take notes with pen and notepad (Kalina & Ziegler, 2015, as cited in Pöchhacker, 2015, p. 411).

In this case, more empirical studies and surveys need to be done to collect interpreters' attitudes towards technology used in SI, in order to figure out whether the negative and reserved attitude towards the employment of technology is a general case for all the people from this profession or just limited to a certain group of interpreters. More importantly, the surveys should aim at helping technology developers figure out what kind of technology is relatively more acceptable to simultaneous interpreters and more likely to be used by them, and what improvements can they make for the existing technologies in SI. What's equally important are psychological experiments designed to analyze and explain interpreters' emotional response to technology-assisted SI, and to truly find out explanations for interpreters' reluctance to use technology. Future researches need to ponder on designing specific measurements and indicators which can help reveal interpreters' discomfort experienced during technology-assisted SI. All these researches are expected to offer findings that will finally contribute to a more interpreter-friendly technology-assisted SI system in the future.

Conclusion

Ever since the progress in electro-technical communication equipment facilitated and contributed to the emergence of SI in conference settings and resulted in conference interpreting as a profession, SI has undergone a long historical period. With a variety of technologies employed in SI, the researches on it have taken on some new features and brought about new inspirations to this industry. Many researches have been done on technology-assisted SI with a variety of research topics ranging from interpreting quality to technological benefits, etc. This paper reviews the historical development of technology-assisted interpreting, summarizes the limitations of the current researches on technology-assisted SI, namely, an imbalance of research topics, and some limitations of researches on SI from technological perspectives, and finally offers three proposals for future research topics in the field of technology-assisted SI: carrying out more empirical studies and surveys to collect interpreters' attitudes towards technology used in SI, conducting psychological experiments to analyze and explain interpreters'

emotional response to technology-assisted SI, and designing specific measurements and indicators to reveal interpreters' discomfort experienced during technology-assisted SI. All these researches are expected to offer findings that will finally contribute to a more interpreter-friendly technology-assisted SI system in the future.

References

Azarmina, P., & Wallace, P. (2005). Remote interpretation in medical encounters: A systematic review. *Journal of Telemedicine and Telecare, 11*(3), 140-145.

Bourlard, H., Dines, J., Magimai-Doss, M., Garner, P. N., Imseng, D., Motlícek, P., … Valente, F. (2011). Current trends in multilingual speech processing. *Sādhanā, 36*(5), 885-915.

Braun, S. (2013). Keep your distance? Remote interpreting in legal proceedings: A critical assessment of a growing practice. *Interpreting, 15*(2), 200-228.

Braun, S., & Taylor, J. L. (2015). *Advances in videoconferencing and interpreting in legal proceedings*. Cambridge: Intersentia.

Fantinuoli, C. (2006). Specialized corpora from the web and term extraction for simultaneous interpreters. In M. Baroni and S. Bernardini (Eds.), *Wacky! Working papers on the web as corpus* (pp. 173-190). Bologna: Gedit.

Fantinuoli, C. (2013). *Interpret bank: Design and implementation of a terminology and knowledge management software for conference interpreters*. Berlin: Epubli.

Haiber, U. (2004). Language recognition systems. In K. U. Carstensen, Ch. Ebert, C. Ebert, S. Jekat, R. Klabunde, and H. Langer (Eds.), *Computer linguistics and language technology: An introduction* (2nd rev. ext. edn.). Heidelberg: Spektrum.

Jekat, S., & Hahn, V. W. (2000). Multilingual Verbmobil-dialogs: Experiments, data collection and data analysis. In W. Wahlster, *Verbmobil: Foundations of speech-to-speech translation* (pp. 575-582). New York: Springer.

Jiang, H. (2013). The interpreter's glossary in simultaneous interpreting: A survey. *Interpreting,* 15(1), 74-93.

Li, X. L., & Wang, M. J. (2018). Construction and research of the teaching model of using automatic speech recognition APP in simultaneous interpreting training course—A case study of VoiceNote as an auxiliary tool. *TEFLE, 40*(1), 12-18.

Lu, X. C. (2020). Distance teaching of interpreting: Delivering simultaneous interpreting courses via video conferencing at GSTI, BFSU. *Chinese Translators Journal, 41*(4), 76-84+191.

Lu, X. C. (2022). Review on 40 years of simultaneous interpreting research in China (1980-2020): Based on the analysis of CSSCI journal papers. *Shanghai Journal of Translators, 37*(2), 31-38.

Moser-Mercer, B. (2005). Remote interpreting: Issues of multi-sensory integration in a multilingual task. *Meta, 50*(2), 727-738.

Pöchhacker, F. (2015). *Introducing interpreting studies*. New York: Routledge.

Roziner, I., & Shlesinger, M. (2010). Much ado about something remote: Stress and performance in remote interpreting. *Interpreting, 12*(2), 214-247.

Sandrelli, A. (2007). Designing CAIT (Computer-Assisted Interpreter Training) tools: Black box. In H. Gerzymisch-Arbogast and S. Nauert (Eds.), *Challenges of multidimensional translation: Proceedings of the Marie Curie Euro Conferences, Saarbrücken, 2-6 May 2005* (pp. 1-18). Retrieved May 29, 2014, from http://www.euroconferences.info/proceedings/2005_Proceedings/2005_Sandrelli_Annalisa.pdf

Sandrelli, A., & de Manuel Jerez, J. (2007). The impact of information and communication technology (ICT) on interpreter training: State-of-the-art and future prospects. *The Interpreter and Translator Trainer, 1*(2), 269-303.

Seeber, K. G. (2012). Multimodal input in simultaneous interpreting: An eye-tracking experiment. In L. N. Zybatow, A. Petrova, and M. Ustaszewski (Eds.), *Translation studies: Old and new types of translation in theory and practice*. *Proceedings of the 1st International Conference TRANSLATA Translation & Interpreting Research: Yesterday-Today-Tomorrow. Innsbruck, May 12-14, 2011.* Berlin: Peter Lang.

Seeber, K. G. (2013). Cognitive load in simultaneous interpreting: Measures and methods. *Target, 25*(1), 18-32.

Stoll, C. (2009). Beyond simultaneous terminology systems. In *Methods of shifting and fixing cognition in the workflow of professional conference interpreters*. Trier: Trier Scientific publishing house.

Sun, H. Q., Li, K. X., & Lu, J. W. (2021). AI-assisted simultaneous interpreting—An experiment and its implications. *TEFLE, 43*(6), 75-80+86+12.

Svoboda, T. (2014). Man and machine: Translation in the era of augmented reality. In W. Baur et al. (Eds.), In the field of tension between man and machine: The future of translators, interpreters and terminologists. *Volume I. Conference proceedings of the 20th FIT World Congress*. BDÜ Weiterbildungs-und Fachverlagsgesellschaft.

Wahlster, W. (2000). *Verbmobil: Foundations of speech-to-speech translation.* New York: Springer.

Wang, X. M., & Wang, B. H. (2021). New trends in the interpreting industry: Mainstream platforms and technologies for remote conference interpreting. *Chinese Translators Journal*, *42*(5), 105-112.

Wright, R. D., & Ward, L. M. (2008). *Orienting of attention*. Oxford: Oxford University Press.

Zhang, A. L., Yang, Z. J., Liu, C. X., & Li, S. (2018). A tentative proposal for translation & interpreting based on human-computer collaboration through developments in artificial intelligence. *TEFLE*, *40*(3), 88-94.