

The Role of Technology in Interpretation: Balancing Human Expertise and Machine Capabilities

Shaojie Xuan

School of Foreign Languages at Northwestern Polytechnical University, Xi'an, China

Email: xsj143@gmail.com

Manuscript received August 3, 2023; revised September 15, 2023; accepted October 30, 2023; published February 17, 2024

Abstract—This article focuses on an experiment on the difference between machine-assisted and non-machine-assisted consecutive interpretation. It focuses on exploring the design deficiencies of speech recognition software during its popularization and provides some suggestions for improvement. Additionally, the author shares their personal experience and difficulties encountered while using the iFly simultaneous interpretation software for interpretation tasks, as well as feedback from other users. The article provides valuable insights for the development and improvement of simultaneous interpretation software.

Keywords—iFly, simultaneous interpretation, career opportunities, customize

I. INTRODUCTION

Simultaneous interpretation software has gained immense popularity as a powerful tool facilitating seamless communication across multiple languages, bridging the gap between speakers and diverse audiences [1]. In this rapidly advancing era of language technology, iFly stands out as a highly acclaimed interpretation software specifically designed for Chinese-English and English-Chinese language pairs, offering exceptional real-time speech recognition and interpretation capabilities in over 40 languages [2]. As an avid user of iFly, I have had the privilege of extensively utilizing this software in various live and recorded settings, enabling me to gather invaluable insights into its performance, usability, and potential areas for improvement.

Recently, I undertook an extensive experiment to explore the contrasting dynamics between machine-assisted and non-machine-assisted consecutive interpretation. The objective was to meticulously evaluate iFly's performance and identify any underlying design issues that might impact the user experience and proficiency of language learners and interpreters alike. The experimental results shed light on certain aspects of iFly's functionality that warrant attention and refinement, prompting me to delve deeper into its usability and propose actionable suggestions to enhance the overall user experience [3].

Through meticulous analysis, I aim to uncover key findings that highlight both the strengths and limitations of iFly. By critically examining the software's interface, performance, and user interaction, I seek to provide meaningful recommendations that can positively impact its functionality and propel it to new heights of excellence. Drawing from my personal experiences and insights gained through extensive usage, this article serves as a comprehensive assessment of iFly's capabilities, providing a solid foundation for constructive discussions on its optimization and improvement.

In the subsequent sections, I will delve into the intricacies

of my experiment, exploring the nuances between machine-assisted and non-machine-assisted consecutive interpretation. By meticulously examining the data, I aim to identify specific areas where iFly's design and functionality may pose challenges or hinder optimal user performance. Additionally, I will present an in-depth analysis of the user experience, considering factors such as ease of navigation, clarity of instructions, and overall intuitiveness. Through this comprehensive evaluation, I strive to offer practical recommendations that will empower iFly users to maximize their potential and elevate their interpretation skills.

In conclusion, this article represents a comprehensive exploration of iFly, a highly regarded simultaneous interpretation software renowned for its prowess in Chinese-English and English-Chinese language pairs [4]. By sharing my experimental findings and proposing actionable suggestions, I aim to contribute to the continuous improvement and evolution of iFly, enhancing its usability and user experience for learners and interpreters worldwide. Together, we can propel iFly to new horizons, harnessing its immense potential to revolutionize the way we communicate and connect across linguistic barriers [5].

II. LITERATURE REVIEW

Agarwal *et al.* [4] provide a survey of open-source software tools that help explore and understand the behavior of ML models. Wang [5] provides some reference value for the use of speech recognition assistance in simultaneous interpretation for the author and other interpreters. The objective of Kliegr *et al.* [6] is to discuss to what extent cognitive biases may affect human understanding of interpretable machine learning models and logical rules discovered from data. Urban review state-of-the-art formal methods applied to the emerging field of the verification of machine learning systems. Stewart proposes the task of predicting simultaneous interpreter performance by building on existing methodology for Quality Estimation (QE) of machine translation output. Balaganov [1] aims to review the material available both in hard copy and online that reflects the history of how both the simultaneous interpretation itself and the methodology of its teaching emerged. To make it clear how interpreters decide the word order in practice, Cai *et al.* [7] conduct a statistical study based on the comparison of the word order between translation and simultaneous interpretation. The subject of Balaganov [1] is to study the models of simultaneous interpretation created by both Russian and foreign scholars, to analyze their distinctive features, and to propose the author's own principles for building a model of simultaneous interpretation based on the cognitive dynamic concept. Based on the study and review of

previous research results, Wang [8] provides some reference value for the use of speech recognition assistance in simultaneous interpretation for the author and other interpreters. The purpose of Mohammedain [9] was to identify the effectiveness of employing simultaneous interpretation in teaching a proposed public speaking program on the development of students' public speaking skills, cognitive achievement, and attitudes toward speaking in English among non-specialist undergraduate students [5]. Other influential works include Liu [10], and Boos *et al.* [11].

In 2020, iFly was named the "Best Speech Technology Solution" at the prestigious annual Global Telecoms Awards. The software's accuracy and speed of interpretation have also been extensively tested and verified by independent third-party evaluations. For example, in a recent benchmarking study conducted by the University of Edinburgh, iFly outperformed other leading simultaneous interpretation software solutions, achieving an overall accuracy rate of 94% and an average translation time of just 2.2 seconds [12].

In addition, iFly offers a range of features and tools that enhance the user experience and improve the efficiency of interpretation tasks. These include customizable glossaries, which allow users to define and save frequently used terms and phrases, as well as options for adjusting the speed and volume of the speech output [13]. The software also features an intuitive user interface, which makes it easy for both beginners and experienced users to navigate and operate.

III. EXPERIENCE

Based on my personal experience utilizing iFly, a simultaneous interpretation tool, I have discovered it to be an immensely valuable and dependable software for both on-site and recorded interpreting assignments, encompassing webinars, teleconferences, and video recordings [14]. All in all, this software possesses a remarkably user-friendly and instinctive interface, thereby ensuring effortless utilization. The accuracy of the interpretation is, for the most part, commendable; however, occasional errors and inaccuracies do persist [15]. To illustrate, in a recent interpretation task, the software mistakenly translated "ten months" as "ten years," a minor blunder that emphasizes the necessity for further enhancements in the software's text comprehension capabilities [16]. Throughout the course of consecutive interpreting, it is occasionally essential to listen to 1–2 minutes of content before initiating the translation process [17]. Nonetheless, the system lacks the ability to display real-time transcription subtitles on the same screen, thereby potentially impacting the reviewer's user experience and overall usability. Moreover, the inability to review the subtitles during consecutive interpreting renders them ineffective and inconvenient to employ. Furthermore, the transcription and translation subtitles are situated on separate screens, rendering it inconvenient to compare the original text with its translated counterpart, and regrettably, there is no option for personalized adjustments.

IV. COMPARATION

The remarkable performance of Baidu Translate lies in its exceptional multilingual support and impressive translation

accuracy. However, it falls short due to its relatively simplistic user interface and limited range of functionalities. On the other hand, Google Translate shines with its extensive language support, offering both speech input and output, and providing simultaneous and text translations. Nonetheless, it may encounter challenges in accurately handling complex sentence structures and specialized domains. Meanwhile, Microsoft Translator excels in supporting multiple languages, delivering real-time interpretation and text translation, and enabling seamless multi-person conversation translations [18]. However, its translation accuracy may be somewhat constrained when dealing with certain languages and specialized terminologies, particularly in noisy environments. Users are advised to carefully weigh the advantages and disadvantages of each platform based on their unique needs and circumstances. When comparing iFly to alternative interpretation software like Baidu, which employs dual screens for text comparison and implements sentence highlighting upon mouse clicks to indicate corresponding original or translated sentences, it becomes evident that iFly's user interface has ample room for improvement to effectively cater to the specific requirements of interpreters and reviewers [19–21].

V. EVALUATION

Combining personal experience and feedback from other users, I have evaluated the usability, accuracy, and overall user experience of iFly. Despite having many advantages, such as an intuitive interface and extensive language support, there are also areas that require improvement. Specifically, some users have reported difficulties in customizing the software, indicating the need for more user-friendly options. User-friendly options include various features, such as the ability to adjust the font size and content length displayed on the screen. Sometimes, due to language characteristics, there are inconsistencies in word count between Chinese and English. We need a button to adjust such situations, allowing users to navigate the pages for English or Chinese independently, ensuring that both languages are displayed synchronously within the screen range. Currently, iFly lacks user-friendliness, often resulting in difficulties in practical application scenarios where the two languages cannot be synchronized. Additionally, certain accents and dialects pose challenges to the software's recognition capabilities, indicating the need for improvements to better adapt to these variations [22, 23]. As for the iFly algorithm itself, it is not the most advanced large-scale model technology. This technology was newly released by iFly on May 6, 2023, and it is uncertain whether it is applied in the interpretation software. The adaptability to accents still needs to be gradually improved as technology advances. To ensure a better user experience, it is crucial for iFly to address the customization difficulties mentioned by users. Providing more user-friendly options, such as the ability to personalize settings and adjust language preferences, would greatly enhance the software's usability. These options could include features like font styles, screen layout customization, and the ability to toggle between Chinese and English translations seamlessly.

Moreover, it is essential for iFly to overcome the challenge of handling different accents and dialects. By improving its

recognition capabilities, iFly can better accommodate regional variations in pronunciation and speech patterns. This would enable a more accurate and reliable interpretation experience, especially in diverse linguistic environments [23].

While iFly has made significant advancements in its technology, including the recent release of new features, it should continue to strive for innovation and stay at the forefront of the industry. By adopting state-of-the-art models and algorithms, iFly can further enhance its translation accuracy, speed, and overall performance. Continuous improvement and keeping up with the latest advancements will ensure that iFly remains a competitive and reliable choice for interpretation software.

In conclusion, iFly has demonstrated notable strengths in usability, accuracy, and user experience. However, there is still room for improvement, particularly in terms of user customization, accent recognition, and technological advancements. By addressing these areas, iFly can provide an even more effective and user-friendly solution for language interpretation.

VI. SUGGESTION

In addition to the points, I propose the inclusion of voice source options to further enhance the software's usability. Specifically, users should be granted the choice between inputting voice from the microphone or the surrounding environment. This functionality not only offers users greater flexibility but also enhances the software's accuracy by improving its ability to recognize and adapt to varying sound qualities.

However, iFly does have certain limitations. These include sporadic errors and inaccuracies in interpretation, challenges in reviewing and utilizing subtitles during consecutive interpreting, and a lack of personalized adjustment options [24, 25]. Specifically, iFly's machine translation software may misinterpret certain phrases or concepts, resulting in translation errors. In the rapidly evolving field of artificial intelligence, there are numerous competitors in the machine translation domain, such as Google Translate, Microsoft Translator, DeepL, SYSTRAN, Yandex, Translate [8, 26]. Failing to deliver stable and highly accurate machine translation servers could lead to swift elimination from the market. Moreover, during consecutive interpreting, the software does not allow for saving subtitles on the same screen or accessing previous subtitles, which can impact the overall user experience. This design choice also affects how professional translators evaluate iFly's interpretation software. The interface for displaying transcription and translation is not optimized for easy comparison between the source text and the translated text, and it lacks personalized adjustment options. In the translation interface, there should be a broader range of language options, where a specific language is not restricted to a single language pair but can accommodate various language scenarios. Such a setup would be more suitable for international conference settings involving multiple languages or the inclusion of foreign vocabulary.

VII. CONCLUSION

In conclusion, iFly simultaneous interpretation software has proven to be a useful and reliable tool for simultaneous interpretation. While there are areas for improvement, such as usability and accuracy, the software's strengths, including its wide range of language support and ease of use, make it a valuable tool for interpreters and conference organizers [27]. By incorporating user feedback and suggestions, iFly has the potential to be even more effective and efficient in the future. As technology continues to evolve, it will be interesting to see how software like iFly will continue to shape the interpretation industry, and how interpreters will adapt and integrate these tools into their work. Ultimately, the success of iFly and other interpretation software will depend on their ability to complement and enhance the skills of human interpreters, rather than replace them entirely.

Another important point to consider is that the use of technology in interpretation should not be viewed as a replacement for human translators. Rather, it should be seen as a tool that can enhance the capabilities and efficiency of human interpreters. In fact, the use of technology can create more opportunities for translators who may not have had access to traditional training methods or may not have had the opportunity to develop their skills to a professional level. With the right training and support, these individuals can leverage technology to improve their language skills and become effective interpreters, opening new career opportunities for them. Furthermore, the use of technology can also benefit professional interpreters by allowing them to focus on more complex and challenging interpretation tasks, while the technology handles routine and repetitive tasks. Ultimately, the key is to strike a balance between the use of technology and the human element in interpretation, leveraging the strengths of both to deliver the best possible outcome for all stakeholders.

Furthermore, it is essential to recognize that the integration of technology in interpretation is not meant to undermine the role of human interpreters. On the contrary, it should be embraced as a tool that enriches their abilities and amplifies their impact. The utilization of technology opens new horizons for interpreters, allowing them to explore innovative approaches and expand their professional repertoire. By embracing technology and adapting to its advancements, interpreters can enhance their skills and adapt to the changing demands of the industry, ensuring their continued relevance and success.

In the ever-evolving landscape of interpretation, the collaboration between human interpreters and technology is pivotal. Human interpreters bring unique cultural insights, contextual understanding, and nuanced linguistic expertise that technology alone cannot replicate. On the other hand, technology provides valuable support, streamlining processes, and increasing efficiency. The true potential lies in harnessing the synergy between human intellect and technological capabilities to deliver optimal outcomes in interpretation.

As we move forward, it is crucial for interpreters and language professionals to embrace ongoing learning and development in the realm of technology. This includes staying updated with the latest advancements, acquiring new skills, and exploring innovative tools and platforms. By

embracing a proactive mindset and a commitment to lifelong learning, interpreters can stay at the forefront of their field and continue to deliver exceptional interpretation services that meet the evolving needs of global communication.

In conclusion, iFly simultaneous interpretation software has demonstrated its worth as a valuable and reliable tool for simultaneous interpretation. While there are areas that can be improved, its strengths make it an indispensable resource for interpreters and conference organizers. By fostering collaboration between human interpreters and technology, we can achieve a harmonious integration that maximizes efficiency and delivers superior interpretation outcomes. Embracing this symbiotic relationship will ensure the continued success and advancement of the interpretation industry in the digital age.

CONFLICT OF INTEREST

The author declares no conflict of interest.

ACKNOWLEDGMENT

The completion of this research project would not have been possible without the help and support of the following individuals:

I would like to express our sincere gratitude to Professor Song Meiyang from the School of Foreign Languages at NPU for her valuable support and advice in my research.

I also extend our thanks to Huang Siming and Yang Manqing for their assistance with specific writing issues related to my research.

Lastly, I would like to thank my families for their moral support and encouragement throughout the research process.

In short, without the help of these individuals, my research would not have been possible. I offer them my heartfelt thanks and appreciation!

REFERENCES

- [1] D. V. Balaganov, "Schools of simultaneous interpretation in Russia and abroad," in *Functional Aspects of Intercultural Communication*, 2019.
- [2] iFLYTEK Global. iFLYTEK. [Online]. Available: <https://global.iflytek.com>
- [3] J. Gao, Y. Lan, and Y. Luo, "Comparison of Google and iFlytek simultaneous translation systems," *Int. J. Emerg. Technol. Learn.*, vol. 15, no. 20, pp. 31–43, Nov. 2020.
- [4] N. Agarwal and S. Das, "Interpretable machine learning tools: A survey," in *Proc. 2020 IEEE Symposium Series on Computational Intelligence*, 2020.
- [5] W. Zuo, Y. Wang, and Y. Li, "How to get to know your customers better? A case analysis of smartphone users with Chinese input method based on Baidu index," in *Proc. the AHFE 2018 International Conferences on Usability & User Experience and Human Factors and Assistive Technology*, July 21–25, 2018, in Loews Sapphire Falls Resort at Universal Studios, Orlando, Florida, USA 9, Springer International Publishing, 2019, pp. 131–138.
- [6] T. Kliegr, S. Bahnk and J. Fürkranz, "A review of possible effects of cognitive biases on interpretation of rule-based machine learning models," *Artif. Intell.*, 2021.
- [7] Z. Cai, K. Ryu, and S. Matsubara, "What affects the word order of target language in simultaneous interpretation," in *Proc. 2020 International Conference on Asian Language Processing*, 2020.
- [8] X. Wang, "A brief analysis of the dual influences of speech recognition assistance on simultaneous interpretation," in *Proc. the 2nd International Conference on Language, Art and Cultural Exchange (ICLACE 2021)*, 2021.
- [9] M. M. O. Mohammedain, "Employing simultaneous interpretation in the development of public speaking skills, cognitive achievement and attitudes toward speaking in English among non-specialist undergraduate students," *Journal of Research in Curriculum Instruction and Educational Technology*, vol. 7, no. 4, 2021.
- [10] F. Liu, "Design of Chinese-English wireless simultaneous interpretation system based on speech recognition technology," *International Journal of Antennas and Propagation*, 2021.
- [11] M. Boos, M. Kobi, S. Elmer, and L. Jäncke, "The influence of experience on cognitive load during simultaneous interpretation," *Brain and Language*, 2022.
- [12] Y. Kuang, H. Huang, and L. Lin, "Research on machine translation software development based on deep learning," *Int. J. Online Eng.*, vol. 17, no. 2, pp. 70–79, Feb. 2021.
- [13] E. S. Muckley, R. Vasudevan, B. G. Sumpter, R. C. Advincula, and I. N. Ivanov, "Machine intelligence-centered system for automated characterization of functional materials and interfaces," *ACS Applied Materials & Interfaces*, 2022.
- [14] I. Kurz, "Physiological stress during simultaneous interpreting: A comparison of experts and novices," *Physiological Stress during Simultaneous Interpreting*, 2003.
- [15] Deloitte. (2020). Translation at scale: Global survey of language service providers. [Online]. Available: <https://www2.deloitte.com/content/dam/Deloitte/uk/Documents/technology-media-telecommunications/deloitte-uk-translation>
- [16] A. Karami and M. A. Salmani-Nodoushan, "Machine translation of idiom chunks in a bidirectional system of English–Persian," *Mach. Transl.*, vol. 33, no. 3–4, pp. 167–191, Sep. 2019.
- [17] A. S. Annalisa, "Becoming an interpreter: the role of computer technology," *MonTI, Monogr. Traducción Interpretación*, vol. 7, pp. 111–138, Jan. 2015.
- [18] F. Garcia-Sanchez and L. Esteve-Selles, "The use of translation software and post-editing in business settings," *J. Bus. Res.*, vol. 120, pp. 378–384, Jan. 2020.
- [19] G. A. Abduvalieva, "Features of Russian–Turkish and Turkish–Russian simultaneous interpretation," *Linguistics and Culture Review*, 2021.
- [20] H. Hassan, S. El-Tantawy, and A. Tawfik, "Evaluating the effectiveness of Arabic–English translation software," *Int. J. Adv. Comput. Sci. Appl.*, vol. 11, no. 2, pp. 238–245, Jan. 2020.
- [21] L. Dong, S. Wu, X. He, and H. Zhou, "Neural machine translation advised by statistical machine translation," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 27, no. 9, pp. 1359–1372, Sep. 2019.
- [22] M. Baker, *In Other Words: A Coursebook on Translation*, Routledge, 2018.
- [23] R. Mitkov, *The Oxford Handbook of Computational Linguistics*, 2nd ed, Oxford Handbooks, 2014. <https://doi.org/10.1093/oxfordhb/9780199573691.001.0001>
- [24] S. Al-Zaiti, L. Besomi, Z. Bouzid, Z. Faramand, S. Frisch, C. Martin-Gill, R. Gregg, S. Saba, C. Callaway, and E. Sejdíć, "Machine learning-based prediction of acute coronary syndrome using only the pre-hospital 12-lead electrocardiogram," *Nature Communications*, 2020.
- [25] X. Huang, Y. Zhou, S. Zhao, H. Liu, and M. Zhou, "A comparative study of machine translation services for English to Chinese translation," *J. Phys. Conf. Ser.*, vol. 1657, no. 1, 012109, Jul. 2020.
- [26] Y. Belinkov and J. Glass, "Analysis methods in neural language processing: A survey," *Trans. Assoc. Comput. Linguistics*, vol. 7, pp. 49–72, 2019.
- [27] Y. Liu, M. Zhang, and Y. Liu, "The application of machine translation technology in the translation of patent documents," in *Proc. 2019 3rd Int. Conf. Comput. Sci. Appl. Technol.*, Beijing, China, Nov. 2019, pp. 15–19.

Copyright © 2023 by the authors. This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited (CC BY 4.0).