Comparing Interface Designs to Improve RSI platforms: Insights from an Experimental Study

Muhammad Ahmed Saeed, Eloy Rodríguez González, Tomasz Korybski, Elena Davitti and Sabine Braun
University of Surrey, Guildford, Surrey GU2 7XH, UK
m.a.saeed@surrey.ac.uk

Abstract. Remote Simultaneous Interpreting (RSI) platforms enable interpreters to provide their services remotely and work from various locations. However, research shows that interpreters perceive interpreting via RSI platforms to be more challenging than on-site interpreting in terms of performance and working conditions [1]. While poor audio quality is a major concern for RSI [2,3], another issue that has been frequently highlighted is the impact of the interpreter’s visual environment on various aspects of RSI. However, this aspect has received little attention in research. The study reported in this article investigates how various visual aids and methods of presenting visual information can aid interpreters and improve their user experience (UX). The study used an experimental design and tested 29 professional conference interpreters on different visual interface options, as well as eliciting their work habits, perceptions and working environments. The findings reveal a notable increase in the frequency of RSI since the beginning of the COVID-19 pandemic. Despite this increase, most participants still preferred on-site work. The predominant platform for RSI among the interpreters sampled was Zoom, which has a minimalist interface that contrasts with interpreter preferences for maximalist, information-rich bespoke RSI interfaces. Overall, the study contributes to supporting the visual needs of interpreters in RSI.

Keywords: Remote Simultaneous Interpreting (RSI), User Experience Questionnaire (UEQ), User Experience (UX)

1 Introduction

In recent decades, technology has intersected with interpreting in various ways, notably in different modalities of distance interpreting and computer-assisted interpreting [2]. In relation to distance interpreting, the uptake prior to the COVID-19 pandemic has been uneven across different interpreting settings. In public-service settings, there has been an increase and diversification of distance interpreting, which further surged during the pandemic [4,5]. However, in conference interpreting, the adoption of distance interpreting was limited before the pandemic despite remarkable technological advancements in remote simultaneous interpreting (RSI). The limited adoption was
attributed to several factors such as poor audio quality and reluctance among conference interpreters to use new, cloud-based communication technologies [2,6].

Initially, RSI was conducted from traditional booth-based environments, such as interpreting hubs. These hubs were connected to remote clients and provided interpreters with the conventional conference interpreting equipment, including the interpreter’s hardware console [6]. By contrast, the new generation of cloud-based simultaneous interpreting delivery platforms which emerged in the years before the pandemic offered interpreters a virtual console and a work environment that simulated traditional interpreting booths [6]. While the new platforms could still be used from within a traditional interpreting booth, they also had the potential to enhance the flexibility of RSI. At the same time, cloud-based RSI also raises several issues. A recent survey [1] indicates that interpreters believe that their performance and working conditions are worse under RSI conditions, corroborating previous findings that interpreters perceive RSI as more challenging than on-site interpreting [2]. While audio quality is highlighted by interpreters as a major factor affecting RSI [2,3,7], another issue that has been consistently highlighted in relation to both booth-based and platform-based RSI is the interpreters’ visual environment [3].

To address this issue, the study reported here, which is part of a larger study focusing on various aspects of RSI, aims to explore the extent to which different aspects of an RSI interface, especially different visual aids and different approaches to presenting visual information, can support interpreters and enhance their user experience (UX). The study drew on an experimental design and tested a cohort of 29 professional conference interpreters in working with different visual interface options, examining the impact of three independent variables on the interpreter’s UX:

- **Type of interface**, with a minimal interface design prioritizing the speaker and featuring hidden controls inspired by Zoom, and a maximal interface design displaying all functions and settings, inspired by bespoke RSI platforms.
- **Interpreter’s view of the speaker**, with a close-up view option displaying only the speaker’s face and a gesture view option showing the speaker’s upper body and hand gestures in addition to their face.
- **Automatic Speech Recognition (ASR)**, with an ASR panel embedded within an RSI interface compared to an interface without ASR.

These visual variables were chosen because they have often been overlooked in RSI research in favour of investigating sound quality effects [2,3,4]. However, previous studies have highlighted the impact of nonverbal visual information, such as hand gestures, lip movement, and body language in interpreting [8]. Other research has emphasized the importance of addressing the interpreters’ sense of presence by thoroughly analyzing their visual requirements [3]. Furthermore, studies on the use integration of ASR in the interpreting workflow have underscored the positive impact that ASR may have for interpreters [9]. In addition to exploring the interpreters’ preferences regarding soft visual aids, we also examined how their attitudes towards RSI were influenced by demographic factors and by their experience in both simultaneous interpretation (SI) and RSI.
This article investigates interpreters’ preferences and usage patterns of various RSI platforms. We also examine the relationship between participants’ most frequently used platforms and their UX, as measured by the User Experience Questionnaire (UEQ) [10] ratings for the tested interface design philosophies, specifically minimalist and maximalist approaches. For brevity, we will not discuss the speaker view and ASR study variables in this article.

2 Methodology

The experimental study reported here used a within-subject design. It was preceded by a focus group with interpreters to inform its design, and supplemented by follow-up interviews with selected interpreters, but the findings from the focus group and interviews are not presented here. For the experimental study, which was conducted between February 2022 and May 2022, we recruited professional interpreters who had at least 400 hours of SI experience and at least 20 hours of RSI experience. The study was conducted online using Qualtrics [11] to guide the participants through the experiment, presenting them with the interpreting tasks, source speeches and questionnaires. In addition, Zoom was used to record the participants’ interpretations. The experiment could be performed within a 24-hour window, according to participants’ availability.

In the part of the study that related to soft visual aids for the interpreter, a two-by-two factorial design was used, involving the simultaneous manipulation of two independent variables. The interpreters were requested to interpret a 25-minute speech, which was divided into two segments. Each segment featured one of two interface designs (minimal/maximal), which were counterbalanced among the interpreters, i.e., some interpreters were initially presented with the minimal design, while others began with the maximal design. Additionally, within each segment, the speaker view (close-up/gesture view) was switched halfway through to ensure that each interpreter experienced both speaker views within each interface design. The order of the different speaker views was also counterbalanced to minimize potential order effects [12]. A publicly available recording of a real-life speech was selected. A crucial selection criterion was that it would display the speaker in the video frame, enabling us to create a contrast between a close-up view and a gesture view.

A pre-experiment questionnaire focused on the interpreters’ RSI-related perceptions, preferences, and work setups. A series of short ‘in-experiment’ questionnaires, administered after each interpreting task, elicited UX ratings based on the short version of the UEQ [13] and qualitative comments from the participants regarding the interfaces presented. Finally, a post-experiment questionnaire elicited the interpreters’ subjective preferences towards the speaker view part of the experiment and any suggestions they had for improving it. We employed a mock RSI interface with two versions, maximal and minimal (see Figure 1) because real-life simultaneous interpreting platforms do not offer users the ability to create multiple distinct versions of the interface as required for our study. For further information on the interface design process, readers can refer to [14].
3 Characteristics of the sample

Out of a total of 35 professional interpreters recruited for the study, 29 successfully completed all of the tasks. Of these 29 participants, 19 identified as female and 10 as male. The majority of participants (89.6%) fell in the range of 30-59, with 8 participants aged 30-39, a further 8 participants aged 40-49, and 10 participants aged 50-59. Two participants were under 30 and one was over 60. Our study’s age and gender distribution shows similarity to other recent studies [1,7], providing contextual information for the current study and facilitating comparison with similar investigations.

3.1 Interpreters’ work experience and preferences

Of the 29 study participants, 23 had more than 1,200 hours of SI experience, while 6 had between 400 and 1,200 hours of SI experience. In terms of RSI experience, 27 participants had over 50 hours, and 2 had between 20-50 hours of RSI experience. The participants interpreted from English into various languages, including Spanish, French, German, Italian, Polish, Latvian, and Cantonese.

Before the pandemic, most interpreters in our sample primarily worked on-site. However, the majority of interpreters in our sample reported an increase in their remote work frequency after the beginning of the pandemic. These findings corroborate other recent findings [1] and suggest that the pandemic has had a significant effect on the work patterns and experiences of interpreters, leading many to rely increasingly on remote work.

The majority of participants in our study preferred on-site work to remote work. Of our 29 participants, 18 either strongly or slightly preferred onsite work, while 7 either strongly or slightly preferred remote work, and 4 had no preference. However, our findings also show that preferences varied to some extent by gender and age. Female participants were twice as likely as male participants to prefer RSI. This may be because RSI, by eliminating the need to travel, makes it easier for women to balance work and family commitments, among other possible factors [15]. Nevertheless, both genders considered RSI to be more difficult and hampering teamwork. In terms of age range, interpreters in the 50-59 age range expressed the
strongest preference for on-site work, despite generally having more experience with RSI than younger interpreters. An interesting contrast was, however, observed for this group. Although 90% of interpreters in this group preferred on-site work, over 50% of their actual workload consisted of RSI.

According to the participants’ qualitative feedback, the major factors contributing to the preference for onsite work included better working conditions, the ability to physically interact with the speaker, and the opportunity to collaborate with colleagues face-to-face. Participants appreciated the opportunity to work in a physical workspace with a well-equipped infrastructure that provided them with an environment conducive to productivity. Conversely, the main reason cited for a preference for remote work was to avoid commuting, which was a significant source of stress and time consumption for participants.

3.2 Interpreters’ platform experience and preferences

As depicted in Figure 2, our pre-experiment questionnaire results show that 28 out of 29 participants had used Zoom with the interpretation function, with 17 indicating it as their most frequently used platform. While 17 participants had used Interactio and Interprefy, only four and two participants, respectively, identified them as their most frequently used platforms. Our findings are consistent with other recent studies [1,7], suggesting widespread use of Zoom with interpreting function for RSI.

All participants also mentioned either having tested or having used more than one platform for their interpreting assignments, allowing the research team to gather valuable insights into comparative user experiences, preferences, and potential strengths and weaknesses of different RSI platforms.

Fig. 2. Interpreters’ platform experience: Blue=number of participants who reported using the platform; orange=number of participants reporting it as their most frequently used platform
The preferred platforms in our sample were Zoom with interpretation function and Interprefy. Both were mentioned as preferred platforms by 11 participants. Interestingly, while participants use Zoom more frequently than bespoke RSI platforms, they almost always prefer the functionality offered by bespoke platforms. They use Zoom mainly because it is imposed by event organizers due to cost or privacy concerns. Some interpreters in our sample appreciated its simplicity and stability, arguing that it has the basic features required for interpreting despite missing some functionalities. Overall, the findings from our pre-experiment questionnaire suggest that the interpreters in our sample have a diverse range of experiences and preferences in relation to RSI and the platforms used for it. We consider this diversity to be a strength of our study as it ensures that the UX testing of visual aspects of an RSI interface is grounded in a variety of experiences and preferences.

4 Interface type: UEQ findings

The User Experience Questionnaire (UEQ) is widely used to evaluate UX across six dimensions, including attractiveness, efficiency, and novelty, and two meta-dimensions: pragmatic and hedonic quality, measuring perceived practical usefulness and user enjoyment respectively, using 26 item pairs rated on a seven-point Likert scale [9]. The UEQ is commonly used to evaluate prototypes in various domains. A short version of the UEQ, consisting of only eight item pairs and focusing on the two meta-dimensions, was developed for studies requiring quick completion and studies such as the present one, where multiple product or prototype versions need evaluation in a single session [10]. We selected the short UEQ for our study and expanded it by including three additional item pairs to measure specific dimensions of the user experience (UX) of RSI. These dimensions are Technical/Human, which evaluates the comprehensibility of the interface design from a human perspective, Alienating/Connective, which assesses whether the interface evokes any sense of alienation or whether it fosters a seamless workflow, and Unprofessional/Professional, which gauges the resemblance of the interface to conventional interpreting tools and RSI platforms.

4.1 Individual UEQ item pair scores

The individual UEQ scores for the two interface types (Figure 3) show that the feature-rich maximal interface outperformed the minimal interface in 9 out of 11 item pairs, particularly in Technical/Human, Boring/Exciting, and Usual/Leading edge. This suggests that users find the maximal interface more human-like, exciting, and cutting-edge. However, both interfaces scored similarly in Complicated/Easy and Confusing/Clear, indicating that participants appreciated the simplicity and clarity of the minimal interface despite their overall preference for the maximal interface.
4.2 Overall UEQ scores

From the overall user ratings (Figure 4), it is clear that the maximal interface outperformed the minimal interface in terms of both usability and joy of use. Specifically, the maximal interface scored almost twice as high as the minimal interface in hedonic quality, which represents user enjoyment, while the minimal interface had a negative score for this meta-dimension. However, the maximal interface only had a slightly better score than the minimal interface in pragmatic quality, which pertains to the perceived practical usefulness of the interface. This suggests that while users find the maximal interface more enjoyable, both interfaces allow them to accomplish tasks effectively.
5 Discussion

This article raises several interesting discussion points. First, there is a notable contrast in the 50-59 age group, where 90% of interpreters preferred onsite work, yet over 50% of their actual workload consisted of RSI. This may suggest that interpreters in this age group perceive their working conditions as particularly adverse. It is possible that interpreters in this age group took on RSI work during the pandemic for financial reasons, i.e., due to a lack of onsite assignments during this time, but also due to increased concerns about travelling and potential exposure to COVID-19, even after the lockdowns, compared to younger interpreters.

Another point that emerges from the experience profiles of participants in all age groups is their preference for onsite work over remote work, stemming from better working conditions, the opportunity to physically interact with speakers, and face-to-face collaboration with colleagues. However, this preference comes with a trade-off in terms of increased stress and time spent travelling, which ultimately results in fewer assignments completed in the same amount of time. This raises the question of whether and under what circumstances a point could be reached where interpreters would consider both SI and RSI to be more or less equivalent.

Our exploration of visual aspects of the RSI interface suggests that visual aspects play a role in the interpreters’ preferences regarding SI vs. RSI. Our comparison of interpreters’ actual platform experiences and the User Experience Questionnaire (UEQ) ratings for each interface reveal a discrepancy between the minimalist interfaces (such as Zoom) that interpreters regularly use and the more comprehensive, bespoke RSI interfaces that they rate higher. Interpreters appear to prefer feature-rich and specialized RSI interfaces while often using more minimalist platforms in their daily work due to factors like cost, accessibility, and market penetration [1,7]. This discrepancy could be one of the factors contributing to the continued dislike of RSI, even though platform-based RSI offers more flexibility than traditional booth-based RSI, such as the ability to work from home.

6 Conclusion

In conclusion, the primary objective of our study is to explore user preferences and experiences related to different interface options, specifically focusing on improving RSI interfaces. We aim to investigate how visual aspects of an RSI interface contribute to the user experience of interpreters and to identify effective ways to display visual information that can enhance the overall UX. In this article, we first examined the preferences and usage patterns of 29 interpreters regarding various RSI platforms. Second, we analyzed the impact of different visual interface versions, a minimalist and a maximalist version, on the interpreters’ user experience, drawing on the short version of the UEQ. The sample characteristics highlight the significant impact of the COVID-19 pandemic on interpreters’ work patterns and experiences, leading to an increased reliance on RSI. However, the majority of interpreters still prefer on-site work, citing better working conditions, the ability to physically interact with speakers,
and the opportunity to collaborate face-to-face with colleagues as key factors contributing to this preference. The UEQ results indicate a preference for feature-rich and bespoke RSI interfaces when objectively evaluated, which contrasts with current practice where interpreters commonly use minimalist interfaces. The maximal interface outperformed the minimal interface in terms of usability and hedonic quality, suggesting that interpreters may benefit from interfaces that prioritize functionality and enjoyment. Our future publications will focus on the other two study variables, namely the speaker view and the inclusion of a source speech transcript created through ASR, as well as on the interview responses gathered from the participants after the experiment.

References

11. Qualtrics. [Internet]. Provo, Utah, USA: Qualtrics; [cited 2023 June 20]. Available from: https://www.qualtrics.com
