Insights Into Challenges Faced by Interpreting Trainees and Their Error Patterns

Reema S. AlDayel
College of Language Sciences, King Saud University, Saudi Arabia

Hind M. Alotaibi∗
College of Language Sciences, King Saud University, Saudi Arabia

Abstract—This study explored the challenges faced by trainees in Arabic and English interpreting by analyzing prevalent errors and investigating their underlying causes. A cohort of 31 female students underwent simultaneous interpretation tests in both languages. The data were used to identify challenges encountered during the interpretation process and analyze participants’ performance using Musa and Al-Maryani’s (2021) eclectic model. The study also examined the impact of L2 level on interpreting output. The findings show that omission is the most common error in interpreting both English and Arabic speech and is often employed as a coping technique for challenges such as speech speed, specialized terminology, structural differences, and numerical information. Hesitation was more pronounced when interpreting English into Arabic, indicating difficulties in directionality and language fluency. Regardless of language proficiency or level, all participants made similar errors when interpreting Arabic and English speech. These findings highlight the importance of developing comprehensive interpreter training that focuses not only on L2 proficiency, but also on strategies for managing the challenges encountered during interpretation, such as speech speed, specialized terminology, and directionality.

Index Terms—interpreting, error analysis, interpreting training, interpreting challenges, Arabic

I. INTRODUCTION

Interpretation is critical in bridging communication gaps across languages and cultures, and its significance in various domains cannot be overstated (Pöchhacker, 2008). However, the field of interpreting is not without challenges, particularly for Arab undergraduate students in the early stages of developing their interpreting skills (Alhiyari, 2013; Al-Jarf, 2022a; Al-Jarf, 2022b).

Given the increasing demand for skilled interpreters in various sectors such as business, diplomacy, and education, the Saudi Arabian context presents a unique setting for studying interpreting challenges. Although there have been significant advancements in the field of interpreting, a research gap remains regarding the specific challenges faced by students and the factors influencing their interpretation performance. Existing studies have predominantly focused on interpreting challenges in general or have focused on professional interpreters, overlooking the unique experiences and needs of Arab interpreters (Alhiyari, 2013; Al-Jarf, 2022a; Al-Jarf, 2022b).

Thus, in this explanatory case study, we investigated the notable challenges faced by students in interpreting and the impact of L2 proficiency on their performance to shed light on the specific challenges that students face in interpreting. Understanding these challenges is essential for developing targeted interventions and strategies to enhance interpreting capabilities. In the next section, we explore the relevant literature on interpreting, interpreter training (IT), and assessment, followed by the methodology section, which describes the participants, data collection methods, and procedures. Next, we present a detailed description of the analysis and results, followed by a discussion of the results and their implications. Finally, this paper concludes with recommendations and directions for future research.

II. LITERATURE REVIEW

The terms “interpreting” and “interpretation” are used interchangeably, and there is a connection to the concept of translation. Pöchhacker (2008) highlights the advantages of interpreting as a form of translation, which allows for a definition based on general translation principles and the unique characteristics of interpretation. According to Kade (1968; as cited in Munday, 2009), interpretation is defined as a form of translation in which the source language (SL) text is presented once and cannot be reviewed, and the target language (TL) text is produced under time pressure with limited chances for correction.

Viezzi (2013) categorized interpretation practices based on mode and setting. Pöchhacker (2008) identifies three primary interpretation modes: simultaneous interpretation (SI), consecutive interpretation (CI), and sight translation (ST). SI involves a real-time spoken translation from SL to TL, whereas, in CI, the interpreter waits for the speaker to...
complete the translation before rendering it. ST is a form of SI that involves the oral translation of written texts. Additionally, Cui and Zhao (2015) introduce the concept of “ad hoc interpreters,” who are untrained individuals performing interpretation tasks, such as family members, bilingual staff, or students.

Viezzi (2013) explains that this setting refers to the context in which interpretation is provided, such as courts, conferences, healthcare, or the media. Cui and Zhao (2015) shed light on the perspective of non-professional interpreters, commonly known as “ad hoc interpreters,” who perform interpretation tasks without formal training. This brief discussion of interpretation practices and the identification of various modes and settings lays the foundation for understanding the challenges that interpreters face.

A. Interpretation Challenges

Interpreter performance is affected by many factors. Jones (1998) noted that interpreters constantly analyze and resynthesize speech and meaning, which is cognitively demanding. Cognitive load is comparable in consecutive interpretation (CI) and simultaneous interpretation (SI), according to Lin et al. (2018). In SI, interpreters use microphones and headphones to deliver real-time speech with a slight delay, which is cognitively demanding and often done by a team that takes turns (Dawrant & Setton, 2016). According to Liang et al. (2017), CI requires more cognitive effort than SI.

Language proficiency and formal training affect interpreting performance. Language proficiency and training affected SI performance and working memory (WM), according to Tzou et al. (2012). Language competency affects interpreting performance and WM, and formal interpreting training may improve language processing. Lin et al. (2018) found that language competency, WM, and directionality affected SI fluency. WM is more important than language for fluent output. Al-Jarf (2022b) found that advanced students were better at English-Arabic interpreting, while beginners showed comparable competence in both directions.

Speech speed and accent pose challenges. Shirinzadeh Bojnourdi et al. (2013) found a correlation between interpreters’ native language-speaking speeds and SI quality. Li (2010) emphasized the impact of fast speech delivery on SI quality. Eraslan (2020) noted that a strong non-native accent can affect interpreter performance. Vogler et al. (2019) suggested the use of intelligent computer-assisted interpreting (CAI) technologies to analyze spoken words and predict keywords that are likely to be left untranslated by an interpreter to minimize translation errors and enhance interpreter performance.

The structural differences between languages and difficulties with numbers, proper names, idioms, and culturally specific terms present additional challenges. Al-Zahran (2021) suggested restructuring utterances to overcome the structural challenges caused by the differences between English and Arabic. Shamuratova (2022) recommends notetaking to help interpret numbers and suggests transcribing proper names and decoding idioms and cultural items. Russo (2010) emphasized the complex cognitive abilities required for interpretation and highlighted the importance of training, including developing general knowledge, comprehension, production of foreign languages, and the coordination of listening and speaking skills.

Understanding the common challenges faced by interpreters can inform the development of training methods that address these issues (Eraslan, 2020; Vogler et al., 2019; Al, 2021; Shamuratova, 2022; Russo, 2010). Several training approaches are described in the following sections.

B. Training Approaches

Gile’s (2009) effort model is prominent in interpreting training and helps interpreters balance mental and cognitive efforts and select optimal strategies. This model categorizes effort into four stages: listening and analysis, memory, production, and coordination. Interpreters begin by receiving and analyzing speech, using their memory to retain information, and employing notetaking as a temporary aid. They then render their recollections and notes into the target language (TL). The effective coordination of these efforts leads to efficient interpretation, allowing trainee interpreters to regulate their energy and avoid mental overload and poor performance.

Cokely (1992) developed a sociolinguistic model to explain how spoken language is interpreted as sign language. This seven-phase model identifies potential miscommunication points during interpretation. Lee (2019) further elaborated on this model by categorizing it into SL- and TL-related phases, providing one-word reminders for each phase, and suggesting related exercises for classroom applications.

The effort models of Cokely and Gile share certain similarities. However, Cokely extends Gile’s first effort by emphasizing the interpreter’s understanding of the speaker’s intended meaning, aiming to bridge the gap between the receiving and producing stages and to prevent miscommunication. Unlike Gile, Cokely focuses on understanding semantic intentions to avoid misinterpretation, rather than coordinating mental efforts. These models provide valuable insights for training future interpreters.

Other approaches include the semantic (Seleskovitch, 1978), interpretive (Lederer, 2010), psychological (Gerver, 1975), information processing (Massaro, 2014), and neurolinguistic approaches (Ahrens, 2011). These models and approaches can be utilized in interpretation classrooms to provide students with the theoretical foundations necessary to support their practices.

Integrating theories in interpreting classrooms can be challenging. However, research and scientific contributions provided instructional strategies, fundamental perspectives on interpretation, and best practices (Gile, 2009).
technology, globalization, and business practices evolve, the translation and interpretation of pedagogy must be adapted (Király, 2003). New technologies and instructional aids such as corpora- and machine-assisted teaching methodologies can be introduced into translation classrooms (Zhu & Wang, 2011). Incorporating technology such as text-to-speech (TTS) software can enhance interpreting training by exposing students to various accents and adjusting the reading speeds. Additionally, student collaboration and feedback can improve interpretation education, and theoretical frameworks such as Gile’s effort model can help students comprehend and address interpreting challenges (Takeda, 2010).

Discourse-based research has identified practical exercises for improving students’ SI skills, such as summarization, paraphrasing, clause relation, and anticipation exercises, to enhance comprehension (Atari, 2018). Incorporating ST into translation and interpretation programs has also shown benefits such as rapid source text analysis, improved note-reading and public-speaking skills, and increased expression flexibility (Li, 2014).

Assessing interpretation quality is essential for evaluating student learning outcomes. Interpretation quality assessment approaches are described in the next section.

C. Interpretation Quality Assessment

The assessment of interpretation pedagogy is crucial for evaluating students’ learning outcomes and ensuring high-quality education (Sawyer, 2004). However, interpretation quality assessment (IQA) is complex and often subjective (Arango-Keeth & Koby, 2003). Researchers have proposed different models and methods for IQA. Pöchhacker (2008) emphasizes the importance of fulfilling the communicative needs and expectations of the intended audience when evaluating interpreters. Communicative competence focuses on understanding grammar and appropriateness in social situations and is commonly used to assess spoken language skills (Savignon, 2017).

A holistic assessment is a widely used method for interpreting evaluations (Lee, 2019; Chen et al., 2022). It provides a general impression of the overall performance but lacks specific feedback on areas of difficulty. Ding (2017) suggested a propositional analysis focusing on the smallest complete meaning units to evaluate interpretation quality. Although useful for examining correctness and errors, this method may overlook other important components of interpretation. Peer assessment can also be employed to actively involve students in the evaluation process and promote learning (Han, 2018; Su, 2019); however, it may lack consistency and impartial judgment.

Several models and rubrics have been developed for IQA (Al-Jarf, 2022b; Clifford, 2002; Goff-Kfouri, 2004; Ibrahim Ahmad Ibrahim & El-Elisy, 2014; Wu, 2011; Al-Maryani & Musa, 2021). Al-Maryani and Musa (2021) proposed an eclectic model based on Riccardi’s (2002) and Na’ja and Abu Mighnim’s (2012) frameworks. Riccardi’s model measures errors at the intertextual and disfluency levels, including pauses and other errors. Na’ja and Abu Mighnim’s model focuses on intratextual errors, particularly syntactic errors. Although these models have been validated and are effective in diagnosing difficulties, their application in IQA requires further exploration.

In conclusion, the complex nature of interpretation poses several challenges, especially for Arab interpreters; however, research addressing these issues is scarce. This study aims to bridge this gap by answering the following research questions:

RQ1: What are the common interpretation errors made by Saudi translation students?

RQ2: What is the impact of L2 proficiency on their interpretation performance?

Through this research, a better understanding of the challenges and factors affecting interpretation quality could be obtained, ultimately informing the development of effective pedagogical strategies for Arab interpreters. The methodological approach used to accomplish the research objectives is outlined below, including descriptions of the participants, materials, data collection instruments, and procedures.

III. MATERIALS AND METHODS

A mixed-methods approach that utilizes both qualitative and quantitative data analyses was adopted. The data collection method involved the following:

A. Simultaneous Interpretation Test (SIT)

An SIT was used to address RQ1 by investigating the interpretation errors made by translation students. The SI mode was selected because it presents significant challenges to students and is commonly used at conferences and international events. The validity and reliability of the test materials were confirmed. Three professional evaluators (PEs), interpretation experts with PhD degrees in translation and 5–10 years of academic experience, validated the materials. They evaluated several types of speech based on language suitability, sound clarity, speech length, and speed. The most suitable speech was selected for the SIT (see Appendix).

For the English speech, a two-minute and 12-s speech by Her Royal Highness Princess Reema Bint Bandar was chosen that consisted of 345 words and had an approximate speech speed of 156 words per minute (WPM). The Arabic speech, delivered by His Royal Highness Prince Muhammed bin Salman, has a duration of two minutes, 188 words, and an approximate speed of 94 WPM. The selected speeches included linguistic elements and challenges relevant to the aims of this study. The accents of the speakers were taken into consideration and speeches delivered by well-known Saudi speakers were chosen to prevent any influence of the speaker’s accent on interpretation output.
B. English Proficiency Test (EPT)

Given that L2 proficiency might affect interpretation performance (Dimitrova & Hyltenstam, 2000), the sample was further stratified using an English proficiency test (EPT). This study used the online Cambridge General English Test (n.d.), which is a shorter version of the Common European Framework of Reference for Languages (CEFR). This test is widely used for assessing and testing 33 languages (Little, 2007); consequently, it can be regarded as an efficient research instrument for a broad assessment of participants’ language proficiency levels. The participants completed the test online by responding to 25 questions. One point was assigned for each question. Participants’ English proficiency levels were automatically determined based on their test scores (Table 1).

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Cambridge General English Test Scores and Corresponding Proficiency Levels (Cambridge Assessment English, n.d.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proficiency level</td>
<td>CEFR level</td>
</tr>
<tr>
<td>Beginner</td>
<td>A1</td>
</tr>
<tr>
<td></td>
<td>A2</td>
</tr>
<tr>
<td>Intermediate</td>
<td>B1</td>
</tr>
<tr>
<td></td>
<td>B2</td>
</tr>
<tr>
<td>Advanced</td>
<td>C1</td>
</tr>
<tr>
<td></td>
<td>C2</td>
</tr>
</tbody>
</table>

C. Participants

A purposeful sample of 31 senior translation undergraduates from the English Language Department of the College of Language Sciences at King Saud University was selected to participate in this study. The participants were female students enrolled in the Simultaneous Interpretation II course, who were native Arabic speakers. They were in their final academic year and had completed three interpretation courses: Consecutive I, Consecutive II, and Simultaneous I. The choice to exclusively include female participants was driven by logistical difficulties in accessing male sections for test administration.

Although the sample size may not be large, case studies often prioritize in-depth analyses and detailed examinations of specific contexts or phenomena. This study focused on understanding the challenges faced by students in interpreting. A sample of 31 participants was considered appropriate for this study. The researchers coordinated with the instructor teaching the Simultaneous Interpretation II course during the third semester of the academic year (2022–2023) to recruit participants.

D. Procedures

The data collection process used in this study comprised four phases. First, approval to conduct the study and access the sample was obtained from the Standing Committee for Scientific Research Ethics at King Saud University at the beginning of the semester.

Second, permission was obtained from the instructor teaching the Simultaneous II course to the enrolled students. This involved attending three sessions in three sections of the course. The sessions took place over three days in the second week of the semester and were conducted in the interpretation laboratory of the College of Language Sciences. The laboratory is equipped with 40 personal computers and microphone headsets and integrates Sanako software, a language learning tool suite that allows students to record their interpretations, and teachers to monitor and assess students’ progress, track attendance, and provide personalized feedback.

During the sessions, the purpose of the study was explained to the students, who were assured that their participation would be voluntary and would not affect their grades. Volunteer participants signed an informed consent form and were instructed to complete an English proficiency test (EPT) on their PCs. After completion, they were briefed on the first speech and interpretation task. English audio was played and the participants began interpreting simultaneously. The recordings were paused to provide a brief description of the Arabic speech. Once the participants were ready, the Arabic speech was played and the recording resumed. After completing the interpretation task, the participants’ recordings were collected and analyzed, as described in the following section.

IV. DATA ANALYSIS AND RESULTS

A. Analysis Framework

To analyze the participants’ interpretation performance, the model of Al-Maryani and Musa (2021) was adopted. This is a validated eclectic model integrating two existing frameworks, specifically the model proposed by Riccardi (2002; as cited in Al-Maryani & Musa, 2021) and the model by Na’ja and Abu Mighnim (2012; as cited in Al-Maryani & Musa, 2021). Riccardi’s model consists of two parts and focuses on measuring errors at different levels: the intertextual level, which includes omission, addition, and substitution errors, and the disfluency level, which encompasses unfilled and filled pauses. On the other hand, Na’ja and Abu Mighnim’s model examines intratextual

1 https://sanako.com/lab-100

© 2024 ACADEMY PUBLICATION
errors made by interpreting students, specifically addressing various syntactic errors. The main syntactic errors in this model pertain to the improper use of gender: singular, dual, and plural (SDP) errors, vernacularism, inflection, and pronouns. This model was selected because it fit well with the analyzed materials, consisting of English and Arabic outputs with different linguistic characteristics (e.g., SDP errors, gender, and other features). The types of errors in the model are listed in Figure 1 below:

<table>
<thead>
<tr>
<th>Error type</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Omission</td>
<td>Refers to information loss because of insufficient comprehension of the ST or difficulties in reformulating the utterance (Riccardi, 2002).</td>
</tr>
<tr>
<td>Substitution</td>
<td>Refers to the mistranslation of an utterance “which is seen as meaning destructive” (Al-Maryani &amp; Musa, 2021).</td>
</tr>
<tr>
<td>Hesitation</td>
<td>This includes all vocalized expressions of hesitation, which have been transcribed as ah, ahim, um regardless of their duration (Tissi, 2000).</td>
</tr>
<tr>
<td>Inflection</td>
<td>Refers to the negative change in the form of a word (i.e., wrong use of form or tense) (Na’ja &amp; Abu-Mighnaim, 2012).</td>
</tr>
<tr>
<td>Correction</td>
<td>Refers to self-correction (i.e., reconstructing) (Al-Maryani &amp; Musa, 2021).</td>
</tr>
<tr>
<td>Addition</td>
<td>Refers to the unnecessary addition of extra-linguistic items that are not included in the ST (Riccardi, 2002).</td>
</tr>
<tr>
<td>Repetition</td>
<td>Repeating a specific linguistic item (a type of hesitation); these occurrences include non-semantic repetitions of a phrase, word, or even part of a word (Tissi, 2000).</td>
</tr>
<tr>
<td>Gender</td>
<td>Refers to the inaccurate interpretation of gender (Na’ja &amp; Abu-Mighnaim, 2012).</td>
</tr>
<tr>
<td>False Start</td>
<td>False starts occur when the speaker interrupts an utterance and begins a new one without having completed it (Tissi, 2000).</td>
</tr>
<tr>
<td>UFPs</td>
<td>Refers to pauses of more than 3 s and not present in the ST (Riccardi, 2002).</td>
</tr>
</tbody>
</table>

Figure 1. Error Type Definitions

B. Analysis Procedure

The analysis involved the following six steps: First, an analysis sheet was designed to facilitate the analytical process and ensure consistency. This sheet included the error types derived from the model, along with abbreviations for each type. It also contained a transcript of source speech divided into translation units (TUs), following Ding’s propositional analysis approach (Ding, 2017). Each TU represents a small unit expressing its complete meaning. Second, the recordings were individually played, transcribed, and analyzed by marking each error with an abbreviation. The third step involved separately calculating the sum of the error types for each participant.

Next, the analysis sheet for each participant was combined with the corresponding consent form to categorize the participants according to their language level (LL). The total error type is calculated separately for each LL. Finally, a random sample of three recordings from each LL, along with analysis sheets containing error-type definitions, was submitted to two professional evaluators (PEs), who were interpretation experts with Ph.D. degrees in translation and 5–10 years of academic experience, for blind analysis to assess inter-rater reliability.

To test the inter-rater reliability, a Cohen’s Kappa Inter-rater reliability test was conducted using an online calculator (IDO statistics, 2023). Cohen’s kappa value was 0.68, indicating “substantial agreement” between the ratings based on the interpretation of Cohen’s kappa results (Landis & Koch, 1977). This ensured the reliability and accuracy of the scoring system used in the analysis.

C. Analysis Results

(a). Interpretation Errors

Figure 2 shows the total number of errors found in English speech interpretation compared to Arabic speech interpretation, categorized according to error type.
In interpreting the English speech, omissions were committed 1118 times, 815 of which were related to the omission of full TUs. Terminologies (e.g., sustainable development, renewable energy, women’s empowerment, and entrepreneurship) were omitted 293 times. In contrast, the numbers included in the speech were only omitted 10 times, representing the lowest occurrence of omission. Substitution occurred 125 times, 16 of which were related to incorrect interpretation of numbers. However, most of the substituted utterances were related to full TUs. Inflection errors occurred 91 times. This refers to the incorrect use of Arabic to render TUs and terminologies. False starts occurred 88 times when the participants interrupted their interpretations and started a new TU. These cases occurred during the interpretation of several TUs. The participants also made 68 vocal hesitations and sounds (e.g., ähm, mm) in several cases, both within one TU and between two TUs.

In addition, various pronouns were rendered incorrectly 45 times, and unnecessary linguistic items were added 30 times. Vernacular, that is, informal Arabic, occurred 19 times when participants rendered several words and phrases (e.g., we, tomorrow, working with, lives, and believe in them) using a Saudi dialect. The participants made 10 self-corrections while interpreting their English speech. Three of these corrections were related to replacing informal phrases with formal ones and one was related to fixing the form of an Arabic word. Another three were used to correct the conveyed meaning and one was used to correct the translated number. Finally, two participants transferred two English words as they were and then made corrections to say them in Arabic. SDP errors were made 10 times, where singular words were rendered plural. In four cases, a few participants repeated the words once or twice as a form of hesitation. Errors related to gender were also observed.

When interpreting Arabic speech, omissions occurred 554 times, 477 of which were related to omitting full TUs, and 37 terms and 40 numbers were omitted. Inflections (the incorrect use of English) occurred 190 times. Incorrect interpretations occurred 133 times, where numbers ranked the highest with 43 incorrect renderings. Terms were also interpreted inaccurately 34 times. The full TUs were substituted four times. In several cases, the participants made 123 vocal hesitation sounds (e.g., ähm, mm). False starts occurred 68 times when the participants interrupted their interpretations and started a new TU. These cases occurred during the interpretation of several TUs. Vernacular use (i.e., the use of informal English) occurred 19 times. Pronoun-related errors occurred 13 times. Self-correction occurred 12 times when some participants attempted to correct the pronunciation or form of words (e.g., pollution, population, and transportation). Unnecessary additions appeared 12 times, whereas a few participants repeated words once or twice in six cases as a form of hesitation. No errors related to the incorrect renditions of singular or multiple items were observed. In addition, no errors related to incorrect renditions of gender were observed. This may be attributed to the fact that the English language system does not have many gender-related forms.

As for the UFPs, the silent pauses that exceeded 3 s were counted based on Riccardi’s (2002) and Cecot’s (2001) definitions and categorizations of UFPs mentioned earlier. Tables 2 and 3 show the UFPs used to interpret English and Arabic speech.
When interpreting the English speech, the counted UFPs indicated that the participants’ UFPs represented 12% of the original speech duration, whereas the deleted TUs represented 54.5% of the total TUs of the original speech. When interpreting Arabic speech, the counted UFPs indicated that the participants’ UFPs represented 12% of the duration of the original speech, whereas the deleted TUs represented 54.5% of the total TUs of the original speech, which is elaborated further in the next section.

(b). Impact of English Proficiency on Error Count

To examine interpretation challenges, it was essential to analyze the impact of L2 proficiency on the participants’ performance. The average total error count of each language group was compared to test the impact of LL on the participants’ interpretations of English speech (Figure 3). Notably, the three groups had comparable levels of committed errors. Furthermore, omission was the most common error type committed by all groups and appeared most frequently among the A2 group.

To investigate the significance of the differences in the participants’ errors based on their LL, a one-way ANOVA was conducted. The results in Table 4 show that the value of the $F$ statistic $= 0.043$ and the $p$-value $= 0.958 > 0.05$, indicating no significant difference in the number of errors committed by the three groups based on their LL at the significance level of $\alpha \leq 0.05$ when interpreting the English speech into Arabic.
Furthermore, the average total error count of each language group was compared to test the impact of LL on the participants’ interpretations of Arabic speech (Figure 4). Notably, the three groups had comparable levels of committed errors. In addition, omission is the most common error type committed by all groups but seems to be less common among the B2 group than the A2 and B1 groups. Hesitation appeared to be more frequent among the A2 participants than the other groups, whereas inflection errors were less frequent among the B1 participants.

To investigate the significance of the differences in the participants’ errors based on their LL, a one-way ANOVA was conducted. The results in Table 5 show that the value of the $F$-statistic = 0.011 and the $p$-value = 0.989 > 0.05, which indicates no significant difference in the number of errors between three groups of participants in the Arabic speech interpretation at the significance level $\alpha \leq 0.05$.

Thus, it may be concluded that specialized terms, numbers, some TUs, language fluency, and interpretation direction were among the main challenges encountered by the test participants. This finding emerged from the analysis of common errors committed by the tested sample. For instance, the omission was heavily adopted to handle some TUs and specialized terminologies in English-Arabic interpreting and vice versa. Additionally, the numbers were repeatedly omitted and substituted when interpreting Arabic into English. Hesitation also appears to be higher when Arabic speech is interpreted in English. Overall, the tested samples exhibited other significant errors related to substitution and inflection. Minor errors were observed, including in SDP, sex, addition, correction, vernacular, and repetition. However, no significant impact on language proficiency was found since participants at all proficiency levels committed similar errors. A detailed discussion of the results is presented in the following section.

V. DISCUSSION

This case study examined students’ interpretation issues and how L2 proficiency affects them. Data showed that omission was the most common error in interpreting Arabic and English speeches, with the latter being higher. Full TUs, terms, and numbers were omitted, indicating that participants were more likely to omit certain elements during interpretation, especially when interpreting English speech. Omission errors were more common when interpreting English speech, suggesting that participants had trouble conveying its content and details. This contradicts Lin et al.
(2018), who found that interpreting students interpret nonnative to native language speech better. Different sentence structures, specialized terminology, and speech speed might be the effecting factors in such a context, which will be explained hereafter. According to the percentages of UFIs and total omitted TUs in English and Arabic speech, the latter were twice as high as the former. Thereby, participants may have relied on omission because they couldn’t keep up with speech speed. This supports Barghout et al. (2015), who found that interpreters make more omissions at higher speeds. The study found that speech speed affected SI, suggesting that omission is a speed-coping technique.

Omissions of full TUs, terms, and numbers reveal the errors’ nature. Omitting full TUs suggests participants struggled to maintain speech coherence and completeness. These findings support Atari (2005), who suggested practical exercises to help Arabic-speaking undergraduates in SI overcome their challenges, including summarization, which can help students deconstruct source speech and separate the main ideas from the details. Students can better understand the speech and avoid omitting important information by focusing on its main point. Paraphrasing helps students identify main and minor themes and rephrase content in their own words, improving comprehension of speech elements. Additionally, understanding how linkers work within an extract or across paragraphs helps interpreters quickly understand the semantic relations of an utterance to avoid interpretation delays by recognizing the connections between parts of a speech.

Atari (2018) suggested a clause relation exercise to help students understand textual relationships between phrases, sentences, and words. This exercise helps students analyze their speech’s structure and organization to improve their interpretation. These practical exercises help interpreter trainees overcome delays and omissions in SI. Technology can be incorporated into these targeted exercises. Al-Jarf (2022c) suggested using TTS software in interpretation training. However, investigating why English speech interpreters make more omission errors than Arabic speech deserves attention.

Omitting terms and numbers suggests the need to learn more specialized terminology and number interpretation techniques. Our study agrees with Dawrant and Setton (2016) that interpreters may not be familiar with technical terms. Interpreters may struggle to accurately interpret specialized terms due to a lack of technical knowledge or experience. Vogler et al. (2019) suggested using intelligent CAI technologies to reduce translation errors and improve interpreter performance. CAI technologies help interpreters predict untranslatable keywords and improve accuracy.

As numbers are important in certain contexts, omitting them can cause significant translation loss. Interpretation trainees must focus their training on numbers interpreting, understanding number systems, and being aware of cultural differences in number representations to accurately interpret and convey numerical information. Practicing note-taking can help in such contexts. Shamuratova (2022) addressed numbers, proper names, idioms, and culture-specific term interpretation issues. Our study found similar results, especially when participants had trouble with numbers and proper names. Notetaking for numbers and transcoding or decoding for proper names and idioms can help interpreters overcome these challenges.

Many inflection and substitution errors were found. Substitution errors often change meanings. Specifically, misinterpreting numbers was higher when translating Arabic speech into English. This suggests that participants struggled to capture the intended meaning or form of certain elements during interpretation. These substitution errors must be investigated and addressed during interpretation training. According to Shamuratova (2022), this could include explicit instructions on how to accurately interpret and convey numerical information and practicing specific procedures for handling numbers, such as notetaking. In addition, most participants misinterpreted the abbreviation at the beginning of the English speech (FII), even though they were informed of its translation before playing the speech but did not write it down. Thus, Cokely’s (1992) interpretation training model advises interpretation instructors to stress notetaking before or during the task.

Verb and noun tenses are typical inflection errors. These errors can make interpreted speech unclear and they suggest participants struggled with grammaticality accuracy in their interpretations. According to Al-Jarf (2022c), interpretation training should focus on linguistic skills. Interpreters must know SL and TL grammar, vocabulary, and language structures to produce accurate interpretations.

The finding that terms, numbers, and full TUs are difficult for participants to interpret and reproduce is important. Technical terms and specialized vocabulary require precise understanding and translation. Targeted vocabulary-building exercises and exposure to specialized texts or materials related to interpreted subject matter can help interpreters learn domain-specific terminology.

Some full TUs were challenging, suggesting that participants struggled to keep the interpreted speech coherent and complete. The meaning and content of longer speech segments may be difficult to capture and reproduce. Training programs should improve note-taking, active listening, and memory retention to address these issues. These findings suggest targeted training and practice to address terms, numbers, and full TU challenges. Interpretation trainees must provide ample opportunities to practice interpreting texts with these elements, using specific exercises and strategies to improve students’ understanding, reproduction, and performance.

The results also show that hesitation was greater when Arabic was interpreted in English. A good command of English was needed because some participants had trouble pronouncing English words like pollution, which delayed them. Since Arabic is the participants’ first language, they may have struggled to recall equivalent English expressions or structure their thoughts to accurately convey their original meanings. Linguistic complexity and differences between
languages may also contribute to higher hesitation. English and Arabic have different grammar, vocabulary, and cultural nuances that make interpretation difficult. According to Al-Rubai (2004) and Al Zahran (2021), difference in English and Arabic word order is challenging. Our findings support this, especially when participants struggled with omissions or restructuring utterances. Interpreters can struggle to convey the meaning and structure of a speech due to language structural differences. To overcome structural challenges, Al Zahran (2021) advises interpretation trainees to focus on restructuring utterances rather than waiting for verbs.

Interpretation training programs should improve participants' TL fluency, provide extensive practice, and offer strategies to improve fluency and spontaneity to address this issue. ST and shadowing exercises (Al-Jarf, 2021) can help participants produce clear interpretations without hesitation. Student’s confidence and self-assurance are also crucial. When they doubt their language skills or make mistakes, interpretation students may hesitate. Thus, a supportive and encouraging learning environment that boosts confidence can reduce hesitation and improve performance.

UFPs were higher in English-to-Arabic interpretation, which is interesting. The higher occurrence of UFPs in this direction of interpretation suggests that the participants may have faced challenges in finding appropriate Arabic words or expressions to convey the intended meaning of their original English speech. This could be due to speech speed, linguistic structures, vocabulary, and cultural references between the languages. Al-Jarf (2022b) explored how directionality affects interpreting competence between beginners and advanced students. Al-Jarf (2022b) found that advanced students were better at English-Arabic interpreting than beginners. The present study found that advanced students made more omissions and UFPs in English-to-Arabic than in the opposite direction, indicating that it is harder. The direction of interpretation can affect interpreters' performance and competence.

Thus, interpretation training programs should improve students’ language fluency and proficiency in TL, provide extensive practice, and offer strategies to improve fluency and spontaneous speech production. As mentioned, exercises that encourage quick thinking, ST, and active listening can reduce UFPs and improve interpretation performance.

Another intriguing finding was that English language proficiency did not affect participants' interpretations of English or Arabic speech. Students with different language skills made similar mistakes. This contradicts Dimitrova and Hyltenstam (2000) and Tzou et al. (2012). However, this study's context and participant characteristics are noteworthy. According to the LL, most participants were “intermediate” learners and the rest were “beginners”. A possible explanation for this finding is that interpretation training or experience strongly influenced participants’ performance, as Tzou et al. (2012) found that formal interpreting training may improve language processing. Therefore, future studies must consider these factors and their potential interactions with language proficiency.

Liang et al. (2017), Dawrant and Setton (2016), and Jones (1998) have noted interpreters' cognitive load. This study found that cognitive load, particularly omission errors, and difficulties with terms, numbers, and full TUs, affected performance. According to Jones, interpreters must constantly analyze and resynthesize speech, which is cognitively demanding. Lin et al. (2018) showed that CI and SI modes have a similar cognitive load. This study's findings, particularly SI's challenges and higher rates of hesitation and UFPs in specific language directions, match Dawrant and Setton's (2016) findings that microphones, headphones, and multichannel equipment in SI mode are challenging. Thus, SI training requires such equipment for real-time interpretation with minimal lag to bring it closer to real-life practice.

Lin et al. (2018) examined language, WM, and SI fluency. Their findings that language competency, WM, and directionality affect interpretation fluency match this study. This study found disfluency and SI quality violations in hesitation and UFPs. Atari's (2018) practical exercises can improve WM's role in interpreting fluency, which is crucial to fluent output.

In summary, the findings obtained from analyzing the SIT results indicate that specialized terms, numbers, speech speed, and structural differences between the two languages pose challenges for interpreting students. Rendering Arabic numbers into English was another challenge, and a higher rate of hesitation was observed when interpreting Arabic into English, owing to differences in the language systems and pronunciation difficulties. Additionally, the participants were not committed to notetaking during interpretation, which resulted in some translational errors. Finally, the findings suggest that the language level of the participants did not play a significant role in the interpretation errors made by the tested group. In other words, regardless of language proficiency or level, all participants made similar errors when interpreting Arabic and English speech.

VI. CONCLUSION

This case study examined the challenges faced by undergraduate students in the field of interpretation and explored the impact of L2 proficiency on their performance. A simultaneous interpretation test was administered to analyze the prevalent errors committed by a cohort of 31 Saudi interpreting trainees. The interpretation performance was analyzed using Al-Maryani and Musa’s (2021) eclectic model. Results revealed that the participants relied heavily on omissions, substitutions, and inflection errors, in addition to hesitation and a significant number of UFPs, to tackle the challenges they encountered in interpreting. These errors can be attributed to several factors such as speech speed, specialized terms, numbers, structural differences, and issues related to pronouncing English words. The findings also revealed that participants’ language level was not a significant factor, as all participants committed similar errors.

This study discusses several pedagogical implications and suggests recommendations for enhancing IT programs. Recommendations involve prioritizing practical training and practice opportunities for ST within specialized translation
courses, providing students with comprehensive glossaries, utilizing equipped labs and technologies, and offering optional training courses that can significantly enhance their interpretation skills. Additionally, focusing on teaching coping strategies as well as linguistic and notetaking skills can help students overcome their challenges and improve their performance.

However, this study has some limitations that might have affected the results’ generalizability. First, it involved a cohort of 31 Saudi female interpreting trainees. The relatively small sample size may limit the generalizability of our findings to a broader population. A larger and more diverse sample could provide a more representative understanding of the challenges undergraduate students face in interpreting. Furthermore, the study focused on female Saudi interpreting trainees, which might have introduced a gender bias, limiting the applicability of the findings to male interpreting trainees. Future studies should investigate the impact of sex on the interpretation challenges among male and female trainees. Finally, long-term assessments or follow-up studies could provide insights into the effectiveness of the recommended pedagogical implications and the impact of training programs on students’ interpretation skills over time.

APPENDIX. TRANSCRIPT OF THE ENGLISH AND ARABIC SPEECH

I want to thank the IFL for the opportunity to join you today and to share with you one of my personal and our country’s highest priorities. Working with supporting empowering our youth. The generation that will carry on the work started by Vision 2030 and making sure it touches the lives of everyone in the Kingdom and across the region. We, as Saudis, are a young nation, 75 percent of our population is under the age of 35, and the course for our young people is not the past, it’s not even the present, it’s the future. The change is in progress. Building a stronger private sector job growth and creation, entrepreneurship, sustainable development, renewable energy, and building a global leading digital infrastructure. Socially, women’s empowerment, greater equality and equity, diversity and inclusiveness. Culturally, placing the Kingdom’s arts and entertainment community on front stage. Our natural preservation efforts sharing our traditions and our heritage with others and the world and the key to our long-term success. The Guardians will ensure that this future society matches our dreams and honors. Our ambition is to create a place for everyone. It’s our young people, their success tomorrow depends on our success today, and if our young people are to be the architects of the future, they need the right tools, and they need us to prioritize their interests. They need us to believe in them, and that means providing opportunity and investing in the future so that their creativity and genius can overcome the challenges that they will face.

We cannot close our eyes, and we must begin the work now. This is what our young people will be able to finish, but we must construct the foundation upon which they can reach the clouds. I firmly believe our priority must be our young people because they are the stewards of the future. Building the future is never complete. Each generation chairs that responsibility of doing its share, empowering and paying the way for the next that’s how the future is built, together.

ACKNOWLEDGEMENTS

This research received grant no. (2023/5) from the Arab Observatory for Translation (an affiliate of ALECSO), which is supported by the Literature, Publishing & Translation Commission in Saudi Arabia.

REFERENCES


Reema S. AlDayel is a researcher with an MA in Translation from the College of Language Sciences at King Saud University in Riyadh. A Skilled translator experienced in simultaneous interpretation and translation. Her research interests include translation and interpretation pedagogy and practice.

Hind M. Alotaibi a Full Professor at the College of Language Sciences at King Saud University in Riyadh. She holds a Ph.D. in Education from The University of Manchester, UK, with a specialization in Computer-Assisted Language Learning. Her research interests include Computer-Assisted Translation, ICT in Education, E-learning, and Mobile-Assisted Learning.