

Comparing Student Self-assessment and Teacher Assessment in Korean-English Consecutive Interpreting

Focus on fidelity and target language

Juyeon Lee
Hankuk University of Foreign Studies
juliajuyeonlee@naver.com

ABSTRACT: Self-assessment in interpreter training has been recognized as an important tool to motivate learners and help them learn to evaluate their own performance. The literature on this topic shows that students' self-assessment ratings positively correlate with trainers' assessments, but there are certain differences between the two groups in terms of interpretation quality assessment. The present study wishes to contribute to research on self-assessment as a learning tool by comparing students' self-assessments with teacher assessments in terms of a set of quality categories and identifying in what areas students need more guidance to draw pedagogical implications. For data collection, 20 first-year graduate students of Korean-English interpretation conducted self-assessment of their sentence-by-sentence consecutive interpretation in both directions. Two broad quality categories were applied, which are fidelity to the source text and target language adequacy, along with the three sub-categories of fidelity which are omissions, misinterpretations, and additions. An experienced interpreter trainer was recruited to perform assessment of the students' consecutive interpretations based on the same quality categories. The results show that the students tended to focus more on target language quality in both directions while the teacher applied stricter criteria in evaluating fidelity of the interpretations. For instance, the study found several instances where students marked parts of their interpretations as target language errors while the trainer marked them as fidelity errors such as omissions and misinterpretations. The results suggest that the students

were not consistent and reliable detectors of errors in their interpretation performance during self-assessment. This may be partly attributable to the fact that the participants were in the second semester of the two-year long graduate program, and in the process of learning how to evaluate the quality of interpretation accurately. Pedagogical implications of the findings are discussed, including the kind of guidance needed for students to learn how to conduct self-assessment more successfully.

KEYWORDS: self-assessment, interpreter training, quality of interpreting, fidelity, target language quality

논문초록: 통역 교육에서 자기평가는 학생들에게 동기부여를 제공하고 자신의 통역을 평가하는 방법을 배울 수 있는 중요한 도구로 여겨진다. 선행연구에 따르면 학생들의 자기평가가 대체로 교수자 평가와 긍정적인 상관관계를 보이지만 자기평가와 교수자 평가는 여러 면에서 차이를 보인다. 이에 본 연구에서는 일련의 통역 품질 기준을 중심으로 학생들의 자기평가와 교수자 평가를 비교하였다. 이를 통해 학생들에게 자기평가에 대한 어떤 종류의 가이드를 제공해야 하는지 조사하고 교육적 함의를 도출하여 통역 자기평가 관련 연구에 기여하고자 한다. 데이터 수집을 위해 통번역대학원 한영과 1학년 학생들이 한영 및 영한 방향으로 문장별 순차통역을 수행한 후 자기평가를 실시하였다. 자기평가 기준은 크게 충실성과 도착어 품질의 두 가지 기준을 사용하였고, 통역품질의 가장 중요한 기준인 충실성은 누락, 오역, 추가의 세 가지 세부 기준으로 평가하였다. 또한 경험이 풍부한 통역 교수자가 학생들의 통역에 대하여 동일한 기준으로 평가하였다. 학생 자기평가와 교수자 평가를 비교한 결과 학생들은 양방향 통역에서 모두 도착어 품질에 더 집중한 반면 교수자는 통역의 충실성 평가에 더 엄격한 기준을 적용한 것으로 드러났다. 예를 들어 학생들이 자기평가에서 도착어 언어 오류라고 표시한 많은 부분들을 교수자는 누락이나 오역 등 충실성 오류라고 평가하였다. 이러한 결과는 학생들이 자기평가를 수행할 때 일관성 있게 오류를 감지하고 발견하지 못한다는 점을 보여준다. 본 연구 참여자들이 아직 1학년 2학기이고 통역 품질 평가 방법을 배우는 과정에 있기 때문에 정확한 자기평가를 수행하기 어렵다고 볼 수 있다. 이러한 결과를 바탕으로 학생들이 효과적으로 자기평가를 수행할 수 있도록 어떤 가이드와 지침이 필요한 지를 포함한 교육적 함의를 논의한다.

핵심어: 자기평가, 통역교육, 통역 품질, 충실성, 도착어 품질

1. Introduction

The goal of graduate programs training interpreting students is to ensure that they acquire the necessary skills to perform as professional interpreters. Students learn basic skills needed to comprehend speeches in the source language and render them in the target language either in a consecutive or a simultaneous interpreting mode. During classes, teachers provide students with feedback and comments on their performance as formative assessment to help learners improve their skills. They learn to provide peer feedback during classes as well as during practice sessions.

When students finish their training and start working as professional interpreters, they are left on their own to evaluate their performance, i.e., engage in self-assessment and continue to search for ways to improve their skills and competencies as interpreters. As a result, it is important for students to learn the techniques of self-assessment during training, which will help them monitor their performance after graduation. As Bartłomiejczyk (2007) notes, self-evaluation should be practiced both by student interpreters and professional interpreters as a useful means of quality control.

Several studies have been undertaken to investigate various aspects of students' self-assessment (Han & Fan, 2020; Han & Riazi, 2018; Lee, 2011; Li, 2018; Shin, 2017; Wu, 2021) based on the assumption that self-assessment is a necessary skill to be practiced by students. Indeed, trainees saw self-assessment as being a useful aspect of the learning process (Araújo, 2019). Previous studies reported that students tend to focus on negative aspects of their performance, and they seem to recognize benefits of conducting self-assessment on their interpretation. In addition, differences between teachers' assessments and students' self-assessments have been investigated through various methods. The present study aims to make a small contribution to the growing body of research on self-assessment in the context of interpreter education by examining differences and similarities between students' self-assessments and teacher assessments in terms of the two quality categories of fidelity and target language quality and three sub-categories of fidelity, namely omissions, misinterpretations, and additions. The patterns emerging from the self-assessments by graduate students of interpretation compared to teacher assessments provide pedagogical implications on how to approach self-assessments for educational purposes.

2. Studies on Self-Assessment in the Context of Interpreter Education

2.1 Assessment in the Context of Interpreter Education

Assessment is an important pedagogical tool that teachers employ to find out what students know and what they can do. In a traditional classroom environment, assessment is typically conducted by teachers, but in a collaborative learning environment, assessment can be conducted by teachers and peers to provide feedback on the learner's skills and performance. In the context of interpreter training, assessment is mainly concerned with quality of interpreting produced by students, which is evaluated for various educational purposes. In discussing how quality of interpreting is not a single measure but a combination of elements, Gile (1995) noted that interpretation quality is "a subjectively weighted sum of a number of components: the fidelity of the target-language speech, the quality of the interpreter's linguistic output, the quality of his/her voice, the prosodic characteristics of his/her delivery, [and] the quality of his/her terminological usage" (p. 151). Considering the long list of components that make up quality in interpretation, one can assume that assessment of interpreting performance is not a simple task.

While quality is multi-faceted and interpretation quality is subject to assessment in the educational setting, Hatim and Mason (1997) make a distinction between assessing quality and assessing performance, whereas quality assessment is concerned with a product, i.e., interpretation output in the target language, and performance assessment is concerned with the process of interpreting. Quality assessment may involve comparing the interpretation with the source speech to see how faithful the rendition is and how appropriate the expressions in the target text are. On the other hand, process-oriented assessment may focus on the skills required to carry out interpreting by using a variety of methods, including a think-aloud protocol and a reflective interview. While there are numerous aspects to assessment, Sawyer (2004) proposes an integrated view of assessment for interpreter education, "an approach that views assessment as providing feedback and guidance to the learner throughout the course of instruction" (p. 93).

Moving on to the types of assessment, it can be categorized in various ways. For instance, Child (2004, as cited in Iaroslavschi, 2011) categorized assessment into four types: pre-task assessment, formative assessment,

diagnostic assessment, and summative assessment. Pre-task assessment is conducted to identify the level of knowledge and skills of students in the beginning of the learning process. Formative assessment is used by the teacher to check how much progress the students have made. Diagnostic assessment is employed during the course to discover the rationale behind difficulties that students experience to help students overcome these difficulties. Finally, summative assessment occurs at the end of the course to measure learning outcomes.

The categorization of assessment by Gipps (1994, as cited in Sawyer, 2004) is similar to the categories described above but includes a type of assessment performed by learners themselves. Gipps (1994) discussed three types of assessment: formative, summative, and ipsative. The first two forms of assessment are almost identical to what Child (2004) described, but ipsative assessment refers to a process in which students evaluate their own performance in comparison to their previous performances. Ipsative assessment not only occurs during the learning process as students but also when they work as professional interpreters. Evaluating one's own performance is a useful method of ensuring quality control, which should be practiced both by trainees and by professionals (Bartlomiejczyk, 2007).

Sawyer (2004) suggests that ipsative assessment needs to be well integrated into the curriculum so that students can fully benefit from self-assessment opportunities to improve their learning. Self-assessment offers several pedagogical benefits by supporting learner-directed and autonomous learning and providing teachers with information about learners in terms of their understanding of interpretation quality.

2.2 Studies on Self-assessment in Interpreting

For the past few decades, there has been major growth in the use of self-assessment in interpreter training. Students are encouraged to self-assess their performance as a means to promote self-directed learning (Shin, 2017) and improve their learning outcomes (Li, 2018). Han and Fan (2020) attribute this trend to three factors. First, there has been a major shift in interpreter education from a traditional, teacher-centered approach to a more student-centered approach to learning (Sawyer, 2004; Setton & Dawrant, 2016). Second, interpreting trainees tend to spend a significant amount of time outside the classroom and practice interpreting on their own or during peer practice

sessions. Self-assessment helps learners develop learner autonomy and develop reflective thinking, thereby supporting their growth as competent interpreters. Third, professional interpreters are required to continue to evaluate the quality of their performance and find ways to improve their skills. Therefore, it is important for students to learn the techniques of self-assessment during training, which can encourage students to engage in a life-long learning process that transcends the classroom (Lee, 2011).

Advantages of using self-assessment for Translation and Interpreting (T&I) education have been discussed in the literature. Li (2018) uses theories on cognitive constructivism in knowledge construction to show the pedagogical benefits of self-assessment in T&I training. Self-assessment is consistent with several pedagogical approaches, including the metacognitive approach to learning, autonomous learning, self-directed learning, lifelong learning, and sustainable assessment (p. 50). Thus, self-assessment is a good tool that supports learners to be reflective of their own learning and engage in the reflective and constructive process of mental construction of knowledge with the help of scaffolding provided by teachers.

Recognizing the importance of self-assessment in interpreter training, several studies have been conducted to examine various aspects of self-assessment (Bartłomiejczyk, 2007; Han & Fan, 2020; Han & Riazi, 2018; Lee, 2017; Lee, 2011; Li, 2018; Wu, 2021). One strand of research is aimed at identifying characteristics of student self-assessment. Bartłomiejczyk (2007) reported in her dissertation study that 84% of the comments in self-assessment by interpreting students were negative comments and only 10% were positive assessments in both directions of English to Polish and Polish to English. To see if there is still a strong tendency towards negative assessment, Bartłomiejczyk (2007) conducted a follow-up study where eighteen interpreting students were asked to interpret a speech from English to Polish, transcribe their interpretations and conduct self-assessment by focusing on both positive and negative aspects of their performance. The self-assessment comments were grouped into product-oriented segments (57%), comments on strategies (22.4%), and others. Product-oriented comments, taking up the largest share of self-assessment, were further classified as relating to faithfulness, coherence, style, lexis, presentation, completeness (i.e., omission), or grammar. Product-oriented comments were also divided into positive (37.1%) and negative comments (56.2%), highlighting a tendency to focus more on negative aspects of their performance during self-assessment. 22.7% of

the positive comments referred to the faithfulness of the interpretation to the source speech while negative comments were mainly related to completeness (23%) and faithfulness (22%). However, issues of presentation were hardly mentioned by the student interpreters, suggesting that guidance is needed to help students pay attention to the presentational aspect of interpretation.

The tendency for trainee interpreters to focus more on negative aspects of their performance was also reported in Wu (2021), where eighteen first-year graduate students of interpretation performed free-style self-assessments of their English-to-Korean consecutive interpretations during a single semester. It was found that most self-assessment comments made by the students were negative comments that pinpointed errors made during interpreting, which may have a negative impact on the learners' self-efficacy. In addition, self-assessment comments provided by the students were largely unstructured and unsystematic at the beginning of the semester but became more structured and organized towards the end of the semester. Students tended to focus on accuracy and target language expressions much more than delivery and note-taking. Comments produced by students during self-assessment constitute a wealth of information that teachers may find useful as they provide clues to the cognitive processes that students go through during interpretation and particular difficulties and issues that they struggle with when trying to produce quality interpretations.

Another research interest pursued in terms of self-assessment was to examine whether trainee interpreters can assess themselves accurately (Han & Riazi, 2018; Li, 2018). Han and Riazi (2018) conducted a longitudinal study to investigate the accuracy level of self-assessments by undergraduate English-Chinese interpretation students and how the accuracy level would change over time for ten weeks. Three scoring criteria were applied to measure self-assessment accuracy: information completeness, fluency of delivery, and target language quality. They found that the self-assessment accuracy generally improved over time for both interpreting directions. The overall accuracy level was higher for English to Chinese interpretation than the other direction. For the other direction, the students tended to over-score the three criteria for Chinese-to-English interpretation at each time point. They were able to self-assess information completeness in the English-to-Chinese direction accurately, but the pattern was reversed in the opposite direction. The authors presume that students may be able to better assess information completeness because they are more competent, confident, and comfortable

in self-assessing their Chinese renditions.

In general, students are capable of conducting self-assessment quite accurately, but certain differences have been reported between students' self-assessments and teacher assessments. In a study conducted by Li (2018), 54 undergraduate translation majors conducted four self-assessments of their sight translation performance using a self-assessment sheet that contained three criteria: consistency (i.e., absence of omissions, additions, and distortions), target language quality, and delivery. The teacher used the same criteria to evaluate the students' performance. The students also filled out a questionnaire at the end of the semester designed to find out their perceptions towards self-assessment. The study found a positive correlation between the students' self-assessments and teacher assessments. The students' self-assessment accuracy improved over time as they conducted repeated self-assessments. Finally, the questionnaire shows that the students view self-assessment as conducive to positive learning outcomes.

A similar study was conducted by Lee (2017) involving fifteen undergraduate students majoring in the interpretation and translation of English and Korean. The students conducted self-assessments of their consecutive interpretation performance in both directions based on the three criteria of fidelity, delivery, and target language. Student self-assessments were then compared with teacher assessments using the same criteria. Overall, the students gave themselves lower scores than the teacher, a finding that is in line with previous findings (Bartłomiejczyk, 2007; Wu, 2021) that interpreting students tended to pay more attention to the negative aspects of their performance than the positive aspects. The students commented more on target language appropriateness for their Korean to English interpretation while they made more comments on fidelity for their English to Korean interpretation. Since English is their L2, they may be more conscious of the correctness of their English rendition, resulting in more comments about target language quality.

Another study on self-assessment was conducted by Han and Fan (2020) on undergraduate students' views and perception of the utility of self-assessment in their learning of English-Chinese interpreting. 38 graduate interpreting students produced reflective observations on their experience of self-assessment throughout a ten-week consecutive interpreting course. The qualitative data analysis revealed four general dimensions of self-assessments: comments, benefits, drawbacks, and suggestions. The study found that the

students generally view self-assessment as beneficial, making them self-aware and reflective of their performance, but self-assessment may also produce differential effects on different groups of students. In other words, some students found self-assessment motivating while others felt discouraged and demotivated when they had to self-assess their interpreting performance. Therefore, the study highlighted the importance of having transparent, explicit, and comprehensible assessment criteria that students can utilize during self-assessment, which helps facilitate benefits of self-assessment and reduces potential negative impact of self-assessment on students.

The previous studies discussed above have informed the design of the present study, but they focused on self-assessments by undergraduate students while the present study was conducted on graduate students. Therefore, a brief review of the studies on self-assessments by graduate students of interpretation is warranted. Lee (2011) made comparisons between self-assessments by 12 graduate students of interpretation and assessments by two teachers collected over two semesters. In each semester, two assessments were conducted on consecutive interpretation so that a total of eight assessment reports were obtained for each student: four self-assessments and four teacher assessments. Both the students and the teachers were asked to produce written comments on interpretations in three assessment categories: meaning, language and delivery. They were also asked to give a grade to each interpretation. The study found a moderately significant correlation between students' and teacher's grades, but students' self-assessments differed from teacher assessments in terms of content. For example, students regarded note-taking, memory, and psychological factors as more important than teachers, and students' self-assessments contained more process-related comments, which can provide a window into students' cognitive processes during interpretation. Shin (2017) and Wu (2021) also reported characteristics of self-assessment at graduate programs of interpretation and translation and benefits of employing this type of assessment in T&I education.

Against this backdrop, the present study aims to contribute to the growing body of literature on self-assessment for interpreter education by examining how student self-assessments may differ from teacher assessments in terms of fidelity to the source text and target language quality. Additional categorization of fidelity into sub-categories of omissions, misinterpretations, and additions is expected to allow more granular comparison between students' self-assessments and teacher assessments.

3. Methodology

3.1 Context of the Study and Participants

To investigate differences between students' self-assessments and teacher assessments on consecutive interpretation between Korean and English in terms of fidelity (omissions, additions, and misinterpretations) and target language quality, the study recruited 20 first-year students at the Graduate School of Interpretation and Translation at Hankuk University of Foreign Studies in Korea. Out of the 20 students, one is male, and the rest are female. Korean is their A language and English is their B language.

In the first semester of the graduate program, students take classes on consecutive interpreting and translation classes in both directions of English to Korean and Korean to English. In the second semester, in addition to the consecutive interpreting classes, they take a course titled Introduction to Simultaneous Interpreting (SI) where they learn the basic and foundational skills needed to perform simultaneous interpretation, for which they train in earnest in the second year of the program. Data were collected during the Introduction to SI classes in the second semester.

As for teacher assessment, to obtain independent assessment data, the study recruited someone who did not teach the above classes but is deemed qualified to evaluate student performances based on extensive experience and knowledge. The teacher who participated in the study has 12 years of experience as a professional conference interpreter and 8 years of teaching experience at the graduate program. She has taught both consecutive and simultaneous interpretation courses in both directions, so she is familiar with the curriculum as well as characteristics of interpreting performance by graduate students.

3.2 Data Collection and Analysis

For data collection, the students were asked to go into simultaneous interpretation booths individually and perform sentence-by-sentence consecutive interpretation without note-taking in both directions, i.e., from English to Korean and from Korean to English. The texts used for data collection contained 20 sentences for each direction (Appendix). The English speech and the Korean speech came from the same conference, which was on

the topic of global infrastructure cooperation, to ensure similar content and difficulty level between the two directions. The topic and basic glossary were provided to the students in advance.

The data collection procedure was as follows. First, the students were each placed in a simultaneous interpretation booth. Since there were only eight booths in the classroom, students were divided into several groups. Right after the author read a sentence, the students interpreted the sentence into the target language immediately without note-taking. The same process was repeated for 20 sentences for English-to-Korean interpretation and another 20 sentences for Korean-to-English interpretation. The students' interpretations were audio-recorded and transcribed by the students themselves.

The assessment categories were informed by several studies. Lee (2017) stated that fidelity, delivery, and language were the most common criteria for interpretation quality assessment suggested in the literature. Han and Riazi (2018) also used three scoring dimensions of information completeness, fluency of delivery, and target language quality. In short, quality of interpretation is judged largely based on how accurate and faithful the renditions are to the original text, how fluent and smooth the delivery is, and how appropriate and natural the renditions are in terms of target language grammar and conventions. In the present study, because assessment was performed on the written transcripts of the interpretation, aspects of delivery (e.g., tone of voice, disfluencies, speed of delivery, intonation, etc.) were excluded from the assessment criteria. As a result, fidelity and language were chosen as the two main assessment categories.

Between the two categories, fidelity should be given more weight due to its importance for successful interpreting. For instance, Moser (1996) examined 201 interviews conducted by 94 interpreters with users at 84 different meetings and reported that the most common expectation of the users was faithfulness to the original speech. Gile (1995) also noted that the most important obligation of an interpreter is to produce a rendition of the speaker's message as faithfully and accurately as possible. Against this backdrop, the fidelity category was further divided into three sub-categories of omission, addition, and misinterpretation, following Gile (1995). Li (2018) also defined content consistency in terms of omissions, additions, and changes of important content. The students in the present study are already familiar with the three sub-categories since they are used in the classroom

when instructors and peers provide feedback on student performance.

As part of self-assessment, students were asked to mark on their transcripts the types of instances that occurred in their interpretation, focused on the three types of fidelity errors (omission, addition, and mistranslation) as well as target language errors. They were also allowed to write comments on certain assessment instances when deemed necessary. The same criteria of assessment were applied by the teacher. Self-assessments by the students were then compared against the assessments by the teacher to examine differences and similarities by assessment category and by interpretation direction and draw pedagogical implications. The results are reported in the following section.

4. Results

The students’ self-assessments and the teacher’s assessments were compared by assessment category and interpreting direction. The results of these comparisons are presented and discussed by interpreting direction.

4.1 The English to Korean Direction

All the instances marked by the students in their self-assessments of the English to Korean interpretations and the assessments by the teacher were tallied in terms of the assessment categories. In all the tables, OM refers to omissions, while MI refers to misinterpretation, AD refers to additions, and TL means target language quality. The instances marked by individual students and those marked by the teachers on the students’ performance are provided in the Appendix, and Table 1 below shows the summary of the tokens identified by the students and the teacher.

Table 1: Assessments of English to Korean interpretations by the students and the teacher

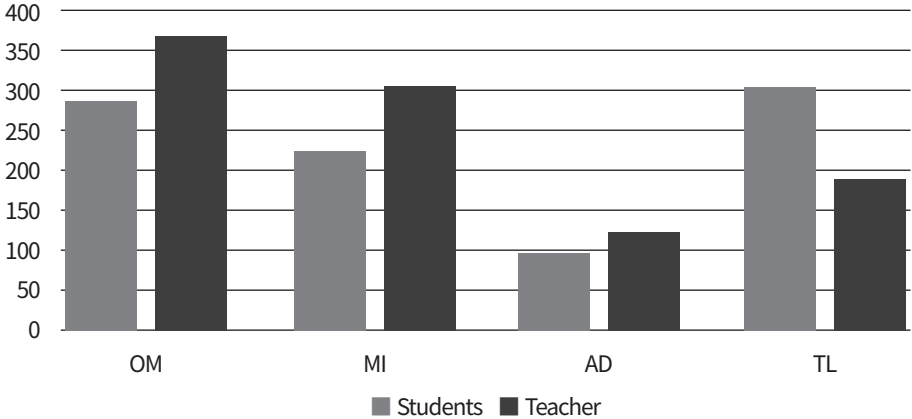
	Fidelity			TL	Total
	OM	MI	AD		
Students	289	227	98	308	922
Teacher	369	305	126	192	819

Table 1 shows that the self-assessments included 614 tokens (66.5%) on fidelity and 308 tokens (33.5%) on target language quality. In the fidelity category, omission errors make up the largest share (47.1%), followed by misinterpretation errors (36.9%), and addition errors (16%). The high frequency of fidelity-related errors may be attributable to the fact that the students did not take notes. Since they had to rely on memory to remember the message and all the details, they may have made omission, misinterpretation, or addition errors when rendering the original message in the target language more frequently than when they interpreted with notes. It is also noteworthy that the students' self-assessments included a large share of tokens on target language errors even though their renditions were in the A language, which suggests that they were either not competent in producing utterances in L1 or unsure of determining appropriateness of target language expressions.

Among the tokens marked by the teacher on the students' English to Korean interpretations, fidelity-related tokens make up the largest share (80.6%) and the target language-related ones account for 19.4%. Under the fidelity category, omission errors were checked most commonly (46.1%), followed by misinterpretations (38.1%), and additions (15.8%). While the students may not have much problem understanding the source speech in L1, they may have trouble remembering details during interpreting because they interpreted without note-taking, which was marked by the teacher in assessing the fidelity of the students' renditions.

When the results of the self-assessments are compared with the teacher's assessments, a clear pattern emerges as shown in Figure 1.

Figure 1: Comparison of assessments of English to Korean interpretations



The comparison between the students' self-assessments and the teacher's assessments shows that the teacher tended to note more fidelity-related errors than the students, indicating that the teacher applied stricter criteria to determine whether interpreted renditions were faithful to the original speech. On the other hand, students tended to pay more attention to target language expressions, showing that the students were overly concerned about target language expressions. The following example is illustrative of these tendencies.

Self-assessment: marked as TL error

Teacher assessment: marked as misinterpretation

ST: I also hope that your concerted efforts will contribute to the realization of the SDGs and the Paris Agreement.

TT: 여러분들의 노고와 고민이 앞으로 파리 협약과 지속가능한 개발목표를 이루는 데 많은 도움이 될 것이라고 생각합니다. (Student 11)

(Literal Translation: I think that your efforts and thoughts will be very helpful in achieving the Paris Agreement and the SDS.)

The underlined part in the source text means that the speaker hopes the ensuing part to take place in the future. The student interpreted this part as a rendition in Korean that denotes "I think", which is marked by the teacher as an instance of misinterpretation but marked by the student as an instance of target language error. The above example shows that the teacher focused on how accurate and faithful the rendition was against the target text and that the student was sensitive to target language adequacy. In fact, the underlined segment marked by the student is perfectly grammatical in the target language. The following are examples that shows how students seem to be sensitive to the appropriateness of target language expressions.

Self-assessment: marked as TL errors

Teacher assessment: not marked as errors

ST: In particular, four of them, namely water and sanitation, energy, cities and communities, and infrastructure are directly related to construction.

TT: 특히 그 중에서도 도시, 물과 위생, 인프라 등 네 개의 지속가능개발목표는

건설업과 직접적으로 연관이 있습니다. (Student 5)

(Literal Translation: In particular, among them, four SDGs such as cities, water and sanitation, and infrastructure are directly related to the construction industry.)

ST: But the impact of construction on the global and national economies is much more profound.

TT: 하지만 건설업이 국내 그리고 전세계 경제에 미치는 영향은 더욱더 거대합니다. (Student 7)

(Literal Translation: However, the impact of the construction industry on the domestic and global economies is much greater.)

In both examples, students marked the underlined parts as indicative of TL errors, but the teacher did not mark them as problematic. In fact, the underlined expressions are appropriate and grammatical in the target language. During self-assessment, students were overly concerned with target language usage to an extent that even appropriate expressions were marked as TL errors.

4.2 The Korean to English Direction

This section reports on how the students assessed their consecutive interpretation from Korean to English as compared to the assessments by the teacher. Table 2 shows the summary of the tokens identified by the students and the teacher during assessment.

Table 2: Assessments of Korean to English interpretations by the students and the teacher

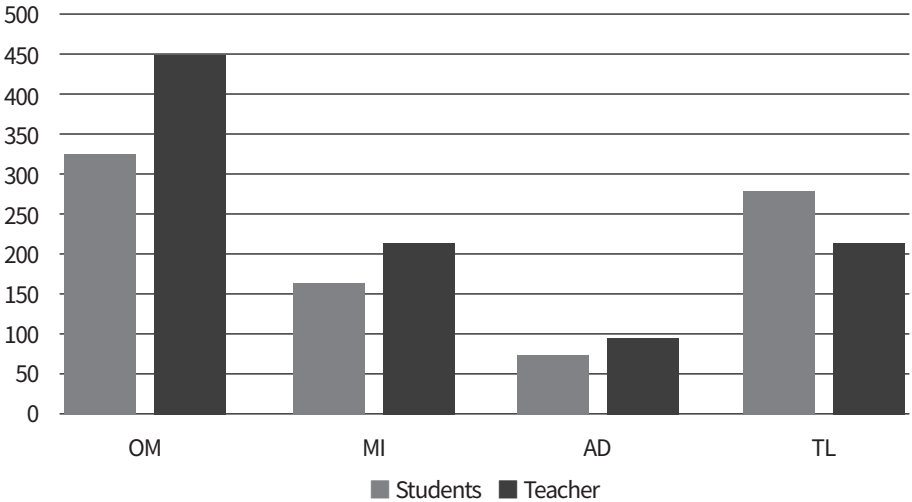
	Fidelity			TL	Total
	OM	MI	AD		
Students	327	168	72	277	844
Teacher	450	213	92	206	961

The self-assessments included 567 tokens (67.2%) on fidelity and 277 tokens (32.8%) on target language quality. Under the fidelity category, omission errors make up the largest share (57.1%), followed by misinterpretation errors (29.7%), and addition errors (12.7%). The distribution of the instances of the assessment categories is similar between the two

directions. Again, fidelity-related errors were marked the most frequently as the students had to rely on memory instead of notes.

In the assessment of the students' English to Korean interpretations by the teacher, fidelity-related tokens make up the largest share (78.6%) and target language-related ones account for 21.4%. In the fidelity category, omission errors were found most frequently (59.6%), followed by misinterpretations (23.2%), and additions (12.2%). The distribution of these instances is similar to the teacher's assessment of the English to Korean interpretations.

Figure 2: Comparison of assessments of Korean to English interpretations



In the Korean to English interpretations, the comparison between self-assessments and teacher assessments revealed a similar pattern as seen in the English to Korean interpretation. In the three sub-categories of fidelity, the students marked more instances than the teacher, but they reported more target language-related issues in their self-assessments compared to the teacher.

As was observed in the English to Korean interpretations, there were many instances in the evaluations of the Korean to English interpretations where target language errors marked by students were not marked as such by the teacher. The following examples show how certain parts of the interpretations were marked by students as target language errors but recognized by the teacher as other instances.

Self-assessment: marked as TL error

Teacher assessment: marked as misinterpretation

ST: 스마트시티는 이 같은 문제를 해결하고 도시의 한정된 자원을 보다 효율적이고 지속적으로 활용할 수 있도록 해 줄 것입니다.

(Literal Translation: Smart City will solve this problem and make it possible to use the city's limited resources in a more efficient and sustainable manner.)

TT: And Smart City is a key to solve this problem in a more sustainable and more meaningful, uh, a more sustainable way. (Student 3)

In the above example, the underlined part in the target language rendition was marked by the student as a target language error but the teacher viewed it as an instance of misinterpretation. It shows that the student was sensitive to target language appropriateness during self-assessment and failed to compare the interpreted rendition with the source text to evaluate fidelity of the interpretation.

Self-assessment: marked as TL error

Teacher assessment: not marked as error

ST: 아세안은 대한민국 국민들이 가장 많이 찾는 지역으로, 연간 방문객 수가 천만 명에 달합니다.

(Literal Translation: ASEAN is the region most frequently visited by people of the Republic of Korea, and the annual visitors are about 10 million.)

TT: About 10 million Korean people visit ASEAN countries every year, which means ASEAN countries are popular tourist destinations for Koreans. (Student 7)

ST: 이를 실현하기 위해서는 아세안과 대한민국의 긴밀한 협력이 무엇보다 중요합니다.

(Literal Translation: To achieve this, close cooperation between ASEAN and Korea is more important than anything else.)

TT: To achieve this community, cooperation between Korea and ASEAN is of utmost importance. (Student 19)

The two examples in the above show how certain expressions were marked by students as target language errors while the teacher considered them as appropriate in the target language. In the first example, the underlined “Korean people” was marked as a target language error in the self-assessment, in which the student noted that the expression “Korean people” would be redundant and should be changed to “Koreans”. This segment was not marked as an error by the teacher. While the student failed to determine accuracy and faithfulness of the entire rendition, too much attention was paid to the details of target language expressions. In the second example, the student marked “to achieve this community” as an instance of target language error but it was not marked as an error by the teacher. The student wrote in the self-assessment that using “achieve” and “community” together did not seem like a proper collocation. When asked to self-assess their performance, students tend to pay attention to details, especially with respect to target language accuracy, when they need to pay more attention to the fidelity and accuracy of their renditions.

5. Conclusion

In the present study, students’ self-assessments and teacher assessments were collected and compared to examine how they may differ in terms of fidelity and target language quality. The results show that students tend to pay more attention to target language quality in their self-assessments than the teacher who applied stricter criteria in determining the fidelity of the interpretations to the source text.

Gile (1995) reported that there was no error that was noticed by all the students and only 20% of the errors were noticed by close to half of the students. Furthermore, about half of the errors were noticed by less than 20% of the students. In short, Gile (1995) concluded that students were not adept at detecting errors reliably. The findings of the present study also seem to suggest that the students were not consistent and reliable detectors of errors in their interpreting performance during self-assessment. This may be partly attributable to the fact that the participants were only in the second semester of the two-year-long graduate program, and in the process of learning how to evaluate the quality of interpretation accurately. It is also plausible to assume that the types of trainer feedback and peer feedback they have received tend

to focus on target language expressions more so than emphasis on producing coherent and faithful renditions.

The study showed that there were certain differences between students' self-assessments and teacher assessments. Yet, as correlation between students' self-assessments and teacher assessments tended to increase with repeated practice over time (Li, 2018), students are expected to master the techniques of self-assessment and assess their performance more accurately with proper guidance and the passage of time. In this regard, the present study may prove useful in that the findings suggest tendencies of student self-assessments that tend to focus on target language quality more than aspects of fidelity. Trainers can thus direct students' attention to accuracy and fidelity of their interpretations and discuss with them what instances constitute accuracy and fidelity errors so that students can better judge and monitor fidelity of their interpretations during self-assessment.

The present study is not without limitations. First, only a small number of trainee interpreters participated in the study, making the results of the study not generalizable to a larger population. Second, the interpretation task employed in the study was sentence-by-sentence consecutive interpretation without note-taking. As such, memory constraints may have induced more fidelity (omission, misinterpretation, and addition) errors than when the students interpreted based on notes. Finally, the self-assessments and teacher assessments were conducted only once. It would be desirable to conduct a longitudinal study where multiple self-assessments and teacher assessments are conducted over a period of a semester or a year to see if students become more capable of conducting accurate and reliable self-assessments with practice and guidance. After all, professional interpreters need to be able to assess the quality of their work and continue to improve their skills and competencies. Therefore, self-assessment skills are important for successful interpreter training.

Appendix

Table 1: Self-assessments by the students of English to Korean interpretations

Participant	Fidelity			TL
	OM	MI	AD	
1	16	10	2	11
2	7	3	1	17
3	16	13	11	16
4	21	10	9	15
5	10	18	3	20
6	9	8	3	5
7	12	5	3	17
8	10	9	2	14
9	9	12	6	15
10	17	24	12	13
11	10	11	8	13
12	10	6	0	17
13	17	12	7	9
14	18	5	0	20
15	12	11	8	33
16	31	16	11	11
17	27	15	6	12
18	12	19	0	26
19	13	18	2	12
20	12	2	4	12
Total	289	227	98	308

Table 2: Assessments by the teacher of English to Korean interpretations

Participant	Fidelity			TL
	OM	MI	AD	
1	21	11	4	12
2	12	16	5	13
3	16	16	5	6
4	16	13	8	8
5	22	20	3	11
6	19	12	15	15

Participant	Fidelity			TL
	OM	MI	AD	
7	23	19	3	11
8	20	23	4	8
9	11	8	6	9
10	17	16	2	11
11	25	21	6	3
12	17	25	8	10
13	13	9	11	9
14	27	8	6	9
15	11	14	8	10
16	31	19	2	9
17	24	14	9	9
18	12	13	6	12
19	17	17	7	10
20	15	11	8	7
Total	369	305	126	192

Table 3: Self-assessments by the students of Korean to English interpretations

Participant	Fidelity			TL
	OM	MI	AD	
1	18	7	2	14
2	14	4	1	14
3	15	7	4	19
4	23	11	4	15
5	17	10	0	16
6	14	10	3	8
7	9	8	3	16
8	7	5	1	11
9	9	10	4	14
10	24	20	5	8
11	13	12	7	20
12	12	2	1	19
13	15	9	2	15
14	17	5	2	10
15	24	5	6	19

Participant	Fidelity			TL
	OM	MI	AD	
16	29	12	3	8
17	17	13	6	6
18	18	12	8	15
19	17	5	6	24
20	15	1	4	6
Total	327	168	72	277

Table 4: Assessments by the teacher of Korean to English interpretations

Participant	Fidelity			TL
	OM	MI	AD	
1	22	11	3	16
2	19	7	3	12
3	21	11	8	12
4	18	9	6	13
5	27	17	2	15
6	28	12	7	5
7	28	10	2	12
8	22	14	4	8
9	21	7	5	12
10	19	9	2	4
11	22	12	3	15
12	26	8	2	10
13	20	6	3	11
14	22	11	8	9
15	25	6	6	9
16	23	9	3	11
17	22	18	7	7
18	20	18	7	3
19	21	9	3	9
20	24	9	8	13
Total	450	213	92	206

Table 5: The source texts used for consecutive interpretation

EK	
1	It is my honor to deliver this keynote speech here at GICC 2019, a major forum on global infrastructure cooperation
2	The construction industry has helped improve our lives and promote the development of communities.
3	Infrastructure has helped realize the basic needs of human life going back millennia.
4	Currently, the scale of the global construction market is around 11 trillion dollars.
5	But the impact of construction on the global and national economies is much more profound.
6	Indeed, infrastructure has a huge effect on business but it is also directly linked to our quality of life.
7	During my ten years as UN Secretary-General, there was no shortage of critical issues confronting humankind.
8	However, I chose to give special attention to sustainable development, including targeted and historical actions needed to combat climate change.
9	The international community agreed that unless we took urgent measures at the local, national and global levels, sustainable development would not be achieved.
10	The 2030 Agenda and the Paris Agreement are concrete promises to future generations.
11	Without a doubt, the processes leading to these agreements were very difficult and challenging, but the implementation of these agreements is even more difficult.
12	The construction industry plays a vital role in international efforts towards realizing the UN's 17 SDGs.
13	In particular, four of them, namely water and sanitation, energy, cities and communities, and infrastructure are directly related to construction.
14	In addition, infrastructure is directly related to the protection of human rights and closing the inequality gap between the rich and the poor.
15	Although the benefits of good infrastructure are enormous, they do come at a price.
16	The construction industry takes up about 40% of the total global energy consumption and it emits 39% of the world's greenhouse gasses.
17	Thus, unless we make efforts for sustainable construction, climate change cannot be holistically mitigated.
18	GICC is a very important forum for project owners, contractors, and financial institutions to gather and discuss cooperation in construction and infrastructure.
19	I am confident that this conference will be successful and productive for all.
20	I also hope that your concerted efforts will contribute to the realization of the SDGs and the Paris Agreement.

KE	
1	안녕하십니까. 각국을 대표해 참석해 주신 아세안 10개 회원국의 인프라 담당 부처 대표님들께 깊은 감사와 환영의 말씀을 드립니다.
2	아울러 이번 회의가 성공적으로 개최될 수 있도록 많은 도움을 주신 대사관 관계자 분들께도 감사드립니다.
3	올해는 아세안과 대한민국이 대화 관계를 맺은 지 30주년이 되는 아주 뜻깊은 해입니다.
4	지난 30년 동안 한국과 아세안은 서로에게 중요한 동반자였습니다.
5	어려운 일이 있으면 서로 돕고 함께 성장하며 끈끈한 우정을 쌓아 왔습니다.
6	아세안은 대한민국 국민들이 가장 많이 찾는 지역으로, 연간 방문객 수가 천만 명에 달합니다.
7	교역과 투자, 건설 분야의 핵심 파트너이기도 합니다.
8	현재 8천 개 이상의 한국 기업이 아세안에 진출해 있고, 무역액은 연간 천6백만 달러에 이릅니다.
9	대한민국은 여기서 한 걸음 더 나아가 아세안과 더욱 가까운 친구가 되고자 합니다.
10	'신남방정책'을 통해 사람, 상생번영, 평화라는 세 개의 원칙 하에 서로에게 선한 영향력을 주는 한-아세안 공동체를 완성할 것입니다.
11	이를 실현하기 위해서는 아세안과 대한민국의 긴밀한 협력이 무엇보다 중요합니다.
12	특히 사람과 사람간의 연결을 위한 인프라 구축은 모든 국가의 경제적 번영을 이끌 것으로 기대됩니다.
13	이는 곧 국민의 삶의 질 향상과 안전 보장으로 이어져 아세안과 대한민국의 평화를 지켜줄 것입니다.
14	아세안과 대한민국이 함께 성장하는 미래 공동체를 완성할 수 있도록 여기 계신 인프라 분야 대표님들께서 힘과 의지를 모아 주시기 바랍니다.
15	올해 인프라 차관회의의 주제는 이 같은 바람을 담아 '스마트시티'로 정했습니다.
16	아세안과 대한민국을 비롯한 전세계의 수많은 도시는 고도성장 과 과밀화, 노후화에 따른 심각한 도전을 맞이하고 있습니다.
17	스마트시티는 이 같은 문제를 해결하고 도시의 한정된 자원을 보다 효율적이고 지속적으로 활용할 수 있도록 해 줄 것입니다.
18	오늘 회의를 통해 지속 가능한 스마트시티의 실현과 한-아세안 미래 공동체 완성을 위한 협력 방안을 논의해 주시기 바랍니다.
19	우리의 노력으로 아세안과 한국의 잠재력이 현실화되고 공동번영의 꽃이 활짝 피어날 것입니다.
20	다시 한 번 대한민국을 찾아주신 귀빈 여러분께 환영의 말씀을 드리며, 한국에서 좋은 추억 많이 만들어가시기를 기원합니다.

Funding

This work was supported by a 2022 Hankuk University of Foreign Studies research grant.

References

- Araújo, Lara D. (2019). Feedback in conference interpreter education: Perspectives of trainers and trainees. *Interpreting*, 21(1), 135-150.
- Bartłomiejczyk, Magdalena. (2007). Interpreting quality as perceived by trainee interpreters: Self-evaluation. *The Interpreter and Translator Trainer*, 1(2), 247-267.
- Child, Dennis. (2004). *Psychology and the Teacher*. Continuum.
- Gile, Daniel. (1995). Fidelity assessment in consecutive interpretation: An experiment. *Target*, 7(1), 151-164.
- Gipps, Caroline. (1994). *Beyond Testing: Towards a Theory of Educational Assessment*. Routledge.
- Han, Chao and Qin Fan. (2020). Using self-assessment as a formative assessment tool in an English-Chinese interpreting course: Student views and perceptions of its utility. *Perspectives: Studies in Translation Theory and Practice*, 28(1), 109-125.
- Han, Chao and Mehdi Riazi. (2018). The accuracy of student self-assessments of English-Chinese bidirectional interpretation: A longitudinal quantitative study. *Assessment and Evaluation in Higher Education*, 43(3), 386-398.
- Hatim, Basil and Ian Mason. (1997). *The Translator as Communicator*. John Benjamins Publishing.
- Iaroslavschi, Maria. (2011). Becoming an interpreter: Assessment and self-assessment in untutored practice sessions. *Cognition, Brain, Behavior: An Interdisciplinary Journal*, 15(2), 229-249.
- Lee, Mi-Gyong. (2017). Suncha tongyeok gyoyukeseo jagipyeonggaeui hyoyongseonge daehan sogo: Hakbu gwajeongeseoeui tongyeok gyoguk saryereul jungsimeuro [Efficacy of self-evaluation as pedagogical tool in undergraduate consecutive interpreting class]. *The Journal of Interpretation and Translation Education*, 15(3), 57-82.
- Lee, Yun-hyang. (2011). Comparing self-assessment and teacher's assessment in interpreter training. *T&I Review*, 1, 87-111.
- Li, Xiangdong. (2018). Self-assessment as 'assessment as learning' in translator and interpreter education: Validity and washback. *The Interpreter and Translator Trainer*, 12(1), 48-67.
- Moser, Peter. (1996). Expectations of users of conference interpretation. *Interpreting*, 1(2), 145-

178.

- Sawyer, David B. (2004). *Fundamental Aspects of Interpreter Education: Curriculum and Assessment*. John Benjamins Publishing.
- Setton, Robin and Andrew Dawrant. (2016). *Conference Interpreting: A Trainer's Guide*. John Benjamins Publishing.
- Shin, Na-Jung. (2017). Jagijudojeok tongyeok hakseup bangbeopeuroseo jagipyeonggaewi gyoyukjeok hyogwae daehan silhaeng yeongu [Action research on pedagogical effects of self-assessment as a self-directed learning tool in interpreter training]. *Interpreting and Translation Studies*, 21(3), 79-107.
- Wu, Kyung Jo. (2021). Tongyeokjeongong haksengduleui jagipyeongga bunseok [Analysis of self-assessment by interpreting students]. *Journal of Learner-Centered Curriculum and Instruction*, 21(9), 667-682.

Professional Profile

Juyeon Lee is Associate Professor at the Graduate School of Interpretation and Translation at Hankuk University of Foreign Studies (HUFS). She currently is the Director of the Interpreting and Translation Research Institute at HUFS. A practicing conference interpreter and a member of the International Association of Conference Interpreters (AIIC), her research interests include interpreter education and processes of interpretation.