Facilitating and Improving Speaking Accuracy through Self-correction and Self-awareness in TOEFL Independent Speaking Tasks

Davood Souri[1]
Ali Merç[2]

Abstract

Assessing the speaking skills of language learners has always been in the agenda considering the fact that speaking correspondingly involves comprehension, and interaction is almost impossible without comprehension in the target language. During the independent speaking tasks of the TOEFL-IBT test, candidates often make grammar mistakes, and since they have limited time for self-correction, they cannot answer the questions fully, and eventually they run out of time finishing the task completely. These mistakes directly affect candidates' speaking negatively, and they often seek ways to make fewer mistakes so that they can focus more on supporting their ideas and fluency. The purpose of this study is to find out if a TOEFL test candidate is likely to use the target language more efficiently in the successive and subsequent trials based on self-correction and self-awareness of errors. For this purpose, a Turkish learner of English preparing to take TOEFL-IBT test in a private English language school was assessed on an 8-step speaking test trial for the 14 speaking questions for the independent speaking tasks of the TOEFL-IBT test. After each trial, the participant's reflections were analyzed through coding and under three themes: lexis, elaboration and structure. The results of this study indicated that, through awareness raising and self-correction, the participant's speaking skills improved in terms of vocabulary range, elaboration, using discourse linkers based on the raters' judgments, however, there was no noticeable improvement regarding sentence structure. The participant also reported that self-correction was an effective way to remember the errors and considered the raters' feedback useful in becoming more aware of certain areas. The interview results revealed that the learner preferred discovery learning rather than rote.

Keywords: speaking accuracy, self-correction, self-awareness, TOEFL IBT testing

Introduction

Interaction and production are the two indispensable components of speaking skills in a foreign language. Considering the fact that speaking involves comprehension, interaction is almost impossible without comprehension in the target language. According to Rivers (1984), adults spend 25 to 30 percent of communication through an oral production of language. Therefore, assessing the speaking skills of language learners has always been in the agenda. Most speaking tests known worldwide, such as IELTS and FCE, are of interactive forms. To illustrate, the examiner asks some questions and the examinee answers them, and then the language produced by the examinee is analyzed based on specific rubrics of the test and then scored. However, when it comes to the TOEFL-IBT independent speaking section, this interaction is absent. In other words, the specific feature that makes the TOEFL-IBT speaking test different from other international tests such as IELTS is that candidates do not speak to a real person, instead they speak to a computer and their

[1]PhD Candidate, Anadolu University, Turkey, e-mail: dsouri53@gmail.com
[2]Assoc. Prof. Dr., Anadolu University, Turkey, e-mail: amerc@anadolu.edu.tr

voices are recorded for future scoring. This lack of real life interaction makes the speaking test a challenge for test takers, and the limited time given to candidates to prepare their answers makes them stressed which eventually may negatively affect their fluency and accuracy.

In the TOEFL-IBT speaking section, there are six tasks: two independent and four integrated ones. The first task is a personal preference task on a familiar topic, and the second one is a personal choice task. For each one, the candidates will hear the speaking task, and they can also see the question, then they need to provide reasons and support their answers. In the second task, candidates are given two choices, and they are required to choose one and give reasons to support their ideas. For each task, they only have 15 seconds to prepare their answers and 45 seconds to speak. Candidates' answers are recorded for future analysis and scoring.

Tasks 3 to 6 are integrated speaking tasks. In task 3, candidates must read a short passage, then they must listen to a conversation on a campus issue, and finally, they must speak based on what they have read and heard and summarize the speakers' opinions on the campus situation in 60 seconds. In task 4, candidates must read a passage on an academic topic, then they must listen to a lecture, and finally, they must speak based on what they have read and heard in 60 seconds. In task 5, candidates only listen to a conversation, there is no reading passage in task 5, on a campus issue and then they must speak about the problem and the solutions mentioned in the conversation and at times explain which solution they think is the best in 60 seconds. In task 6, candidates only listen to a lecture on an academic topic, there is no reading passage, and they must summarize and mention the key point of the lecture in 60 seconds. Candidates are given 30 seconds to prepare their response for tasks 3 and 4 and 20 seconds for tasks 5 and 6.

For scoring the speaking section in the TOEFL-IBT, raters score the candidates responses on a 0 to 4 scale. Raters will be looking for three main categories while scoring candidates' responses. The first category is related to fluency and phonology. Raters look for how fluent and clear the candidates' responses are. In this part, raters take clear and acceptable pronunciation, intonation as well as natural pace into account. The second category revolves around the structure and lexis of English language, and the third category is mainly about pragmatics. The raters look for how fully the candidate's response is, in other words, how clearly he/she has expressed and connected his ideas in a way that is easy to follow. The following paragraphs discuss the main issues regarding speaking skills namely assessing speaking skills, error correction and metacognition.

Assessing speaking skills

Assessing speaking skills has always been a controversial and complex issue among English teachers and test raters since there are no clear-cut rubrics for speaking assessment, and EFL learners have idiosyncratic characteristics, and when it comes to computer based language testing it becomes more complicated. Brown (2001), states that it is not simple and easy to determine valid, reliable, and practical methods to assess oral proficiency traits. It should also be noted that the speaking raters' internal consistency and reliability make speaking assessment more complicated. Brown (2001) believes that a good testing oral proficiency should have four main characteristics. First, it should be a live performance rather than a recorded one. Second, the test-taker relationship should be of one-on-one. Third, a careful specification of the task should be accomplished during the test, and fourth, a descriptive scoring rubric should be provided. There are several components such as grammar, vocabulary, pronunciation, intonation, fluency, and appropriateness of the language, to be taken into account in speaking tests and such components make assessment

extremely complicated since there are strong disagreements among ELT teachers and test raters on how to evaluate and assess the mentioned components. Madsen (1985), claims that many teachers are not sure what criteria to select to assess oral communication and at times they even do not try to evaluate their students' speaking skills. There is another area of controversy in terms of assessing English language learners 'speaking proficiency. Some claim that there are differences in terms of rating learners' speaking proficiency between raters who are native speakers of English and nonnative speakers of English; however, studies show similarities and almost no significant difference when it comes to rating learners' speaking proficiency between native and nonnative speakers of English. Zhang and Elder's (2014) conducted a research in which they explored the judgments of native and nonnative English speaking teachers on oral proficiency based on the College English Test-Spoken English Test. The results of their study showed that although both native and nonnative English speaking teachers differed in terms of their approaches to rating, they were similar when it came to overall rating of both groups. Ekmekçi (2016) also conducted a similar research comparing native and non-native English language teachers' judgments on oral proficiency of 80 EFL students attending preparatory classes. The results revealed that raters had high inter-rater reliability coefficients and there was no statistically significant difference in overall scores given by the two groups of raters.

Error correction
Making errors is an indispensable part of a learner's learning process and correction of such errors by teachers is a very complicated issue since there are disagreements among teachers on how errors and when errors should be treated. Some teachers as well as teacher trainers believe that teachers should only correct their students' errors when they cannot correct themselves. However, language learners expect their teachers to correct their errors that is why error correction is still a controversial issue. In a research conducted by Makino (1993) on 62 EFL students, he tried to find out whether teacher hints could help their students correct their own errors made in their written compositions. The results revealed that students' linguistic competence was activated by the cues given to them by their teacher to correct their own mistakes. Another important issue regarding error correction amongst English teachers is that their students sometimes avoid structures which they are uncertain of, so English teachers are not fully aware of their students' errors which in turn makes error correction a complicated issue. Chaudron (1977) came up with a new term which he defined as "corrective feedback" which laid emphasis on teachers' reactions to errors students make. He believed that due to the reaction of teachers on learners' errors, they tended to transform and improve their utterances. Based on the feedback learners receive, they tend to respond differently. Corrective feedback is of different types such as explicit correction, recasts, elicitation, and repetition of error, clarification, and metalinguistic feedback. Some educators such as Truscott (1996) believe that oral correction of grammar mistakes interferes with meaning making; on the contrary, some other educators such as Ming-chu and Hung- chun (2009) claim that learners whose grammar mistakes are corrected orally are less likely to make the same mistakes in real life. It should be noted that students and teachers' opinions on error correction vary considerably, and when both teachers and students' expectations move towards the same point, teachers are more likely to guide their students to successful language learning. If their expectations diverge, it may end up with dissatisfaction (MacIntyre & Gardner, 1994). One of the most important issues regarding error correction is the time at which correction is provided. The literature classifies error

correction in two broad categories: immediate and delayed feedback. Behavioristic psychologists such as Skinner strongly supported immediate feedback. In behavioristic view, error correction is considered as reinforcement for correct responses. However, some other scholars and linguistics strongly disagree with immediate error correction. They claim that errors if corrected later, can be better recalled and retained. According to Lemley (2005) instant error correction interferes with learners' abilities to better understand the correct answer. If learners' errors are corrected later, learners will have enough time to forget the initial incorrect responses and will be able to replace the correct ones with the incorrect ones. However, Hendrickson (1978) believes that one cannot say what time is the most appropriate and the best time for error correction. According to a research conducted by Farahani and Salajegheh (2015) 460 English teachers and EFL students completed Fukuda's (2004) questionnaire on error correction. The results revealed that teachers preferred immediate error correction while the students had a strong preference for delayed one.

Metacognition and awareness raising

When it comes to speaking, the important role of metacognition becomes more obvious. Canale and Swain (1980) in their communicative competence framework claim that there are certain abilities through which one can achieve speaking proficiency. These abilities include grammatical competence (morphology, semantics, syntax, and phonology), discourse competence (sociocultural rules and discourse), and strategic competence (verbal and non-verbal communication strategies). The grammatical and strategic competence types are directly related to metacognition. Flavell (1979) was the first person who presented a formal model of metacognition. He defined metacognition as someone's knowledge which concerns his/her cognitive processes and products. Wenden (2001) related metacognition to long term memory which contains learners' awareness and knowledge about the process of learning. Regarding speaking skills, the steps EFL learners take and the decisions they make directly affect their speaking. This lays emphasis on the fact that the more learners are aware of learning the better decisions they may make in speaking. Vandergrift (1999) believes metacognitive strategies which include planning, monitoring and evaluating help learners to regulate and direct their learning process. It should be noted that metacognition falls into two main categories which should not be regarded as similar ones. The first category is metacognitive knowledge. According to Brown et al. (1983), this category is related to information which is acquired by learners about their learning. However, the second category, the metacognitive strategies, refers to general skills through which learners plan, monitor, and evaluate their learning. It is wise to say that without awareness and knowledge, learners cannot come up with efficient and effective strategies. Ellis and Sinclair (1989) also claim that learners' awareness of their language learning process helps them acquire learning strategies. Dormer (2013) divides awareness into two main categories: motivation and input. According to this categorization, there are three subcategories for each main category. For the motivation category, first, learners should be aware of their language deficiencies, then they should be made aware of the significance of such deficiencies and finally, they should come to this conclusion that change is possible. For the input category, initially, learners should be aware of the structures and meaning connections, then they need to find out their personal needs in the use of language and why they are using the language and finally, they need to be made aware of details in receptive language.

Statement of the problem and the research question

During the independent speaking tasks of the TOEFL-IBT test, candidates often make grammar mistakes, and since they have limited time for self-correction, they cannot answer the questions fully, and eventually they run out of time to finish the task completely. One of the main concerns of the learners preparing for the test is the grammar mistakes that make them anxious, directly disaffecting their speaking performance. They, then, often seek ways to make less mistakes so that they can focus more on supporting their ideas and fluency. Hendrickson (1978) believes that error correction can help EFL/ESL students improve their proficiency level. He further points out that only those errors which are made frequently and the ones which stigmatize English learners as well as inhibit communication should be corrected. Considering the related literature, the following research question was formed:

Is a TOEFL test candidate likely to use the target language more correctly in the successive and subsequent trials based on his self-correction and self-awareness of the errors?

Methodology

The research design: The present study was designed in the form of an action research as this type of a design looks for an organized route followed by educational practitioners to collect evidence about and afterwards develop their teaching and their students' learning (Mills, 2000). As Creswell (2012) points out, this research study aims to address a specific and real problem within the teaching/learning context specified in the present inquiry and to examine the actual practices of the researchers. The following sections describe the methods used in the present investigation.2.2.

The participant: The subject of this study was a male Turkish learner of English (27 years old) preparing to take TOEFL-IBT test in a private English language school in Istanbul. He had completed his upper-intermediate level in the same school, and he was highly motivated to take a high score in the test. Gaining university admission for master's degree and improving his career were the two main reasons why he had decided to take the test. The school had given permission to the author of this paper to conduct the research.

Data collection and instruments: The present study made use of a qualitative design. The reason for this study was to evaluate the extent to which the test taker's self-correction and self-awareness of his errors would affect the improvement and development of his speaking skills in preparation for the TOEFL-IBT independent speaking tasks. The instruments used in this study included 14 speaking questions for the independent speaking tasks of the TOEFL-IBT test. The first question in each task was a personal preference question on a familiar topic, and the second question was a personal choice one. For each question, the participant heard the speaking task and at the same time, he could also see the question. Then he had 15 seconds to prepare his response and 45 seconds to speak. It should be noted that Cambridge Preparation for the TOEFL Test book provides a CD on which there are seven complete mock tests. The speaking questions were all from the CD. The following steps were taken each time the participant spoke on the task for seven consecutive weeks.

Step 1: The test taker heard and saw the first independent speaking question of the TOEFL-IBT test on the computer monitor, then he prepared his response within actual time limit (15 seconds for preparation and 45 seconds to speak) and then he spoke. After that, he heard and read the second

independent speaking question and did the same process. His speaking was recorded for both questions.

Step 2: The researcher sent the recorded speaking tasks file as an attachment to the participant's e-mail right after the speaking part was over.

Step 3: The participant corrected his mistakes at home later and emailed the actual transcribed speaking and its correction as an attachment to the researcher.

Step 4: Two native speakers of English (one American and one New Zealander) rated the test taker's speaking based on the TOEFL-IBT speaking rubrics and then gave him a score between 0 and 4. They both listened to the learner's speaking recording twice each time.

Step 5: Before each trial, the researcher provided feedback and helped the participant correct his errors if and only if he had not noticed his errors. Long (1977) draws a distinction between error correction and error feedback. He believes teachers' feedback is error detection rather than error correction which eventually promotes correction, and it is not correction in itself. The participant was also asked to read the rater's comments on his speaking parts. In this step, the researchers explained some terms written by the raters such as elaboration, compound, complex structures, and intelligibility. The researcher also tried to help the student add additional grammatical points such as, linkers and conjunctions, based on the raters' comments, through eliciting and awareness raising rather than direct instruction.

Step 6: The participant was asked to write his reflection on each trial.

Step 7: These steps were repeated seven times; however, each time the student spoke about new questions.

Step 8: The researcher compared the raters' scores through the whole process to find out to what extent self-correction and self-awareness of the student had helped the participant improve his speaking for the TOEFL-IBT test in the end.

Data analysis procedure: The qualitative data regarding the participant's reflections were organized by the researcher initially and later were analyzed through coding. By the use of the codes, the researcher identified three themes in the participant's reflections namely lexis, elaboration, and structure. Under the lexis theme, there was only one subtheme, namely limited range of vocabulary. However, under the elaboration theme, there were two subthemes as examples and linkers. The third theme was sentence structure with repetition of the same structure and using simple structures as subthemes.

Findings

The participant's reflections: The following themes were identified and labeled based on the participant's reflections.

Lexis: The participant mentioned in his reflections that he needed more vocabulary for the task.

I couldn't use phrase and enough vocabulary.

I should be relaxed and think of more different examples and words.

According to Harris (1974) speaking as a complex skill requires five important components one of which is vocabulary. He believes these five components are used simultaneously which makes speaking a challenging skill to master. The researcher helped the participant be more aware of using examples to support his ideas through eliciting. He asked "*Is it enough just to answer the question?*" and when the answer was '*No*', the researcher asked the participant how he would

support his idea/s. At this point the participant mentioned that he could support his ideas by using examples. Although the participant may have used examples to support his ideas in the past, he was not aware of the importance of examples and how using examples could help him use better vocabulary while answering the question. Moreover, the researcher asked the participant if repeating the words would help him use a wide range of vocabulary. This question made the participant more aware of the fact that he needed a wide range of new vocabulary rather than repeating the same words. This awareness became more apparent in his further reflections which the researcher labelled it repetition as a new theme.

Elaboration: The participant mentioned several times in his reflections that he needed to elaborate his speaking more through using examples and discourse linkers.

Examples:

I am not able to elaborate my speech and give more examples.

I didn't give more examples and it was the reason I did not elaborate my speech.

I should be relaxed and think of more different examples and words.

I need to give more examples and make myself clearer.

Linkers:

I don't use linkers when I speak.

I also need to practice on linking words.

In 5th trail I am able to use different linkers.

In the 6th trial I have some irrelevant linkers.

My linker should cause elaboration of text.

In the last trial, I wrote down some linkers and while I was speaking, I could use them.

The raters' feedback helped the participant be more aware of the lack of discourse linkers and examples in his speaking to better elaborate on his reasons.

Sentence Structure: The participant mentioned that he repeated himself in terms of both vocabulary and grammar as follows:

Also, I repeated myself. I used the same sentences.

I am repeating my words.

I repeated the modal should frequently

Analysis of the trials

Through a detailed analysis of all the seven trials, three themes were identified: vocabulary range, cohesive devices, structure.

The consistency of the raters' scores: The table below provides a clearer visual comparison between the two raters' scores. As can be clearly seen, it shows the scores given by the raters after each speaking trial. For similar scores the researcher gave 1 and for the different one 0. Then the total number of similar scores (6) was divided by the total number of trials (7) to find out how consistent the raters' scores were. 6/7: 0.85.

Table 1. Scores given by the raters after each speaking trial

	Trial 1	Trial 2	Trial 3	Trial 4	Trial 5	Trial 6	Trial 7
First rater	2/4	2/4	3/4	3/4	3/4	3/4	3/4
Second rater	2/4	2/4	2/4	3/4	3/4	3/4	3/4
Similar: 1 Different: 0	1	1	0	1	1	1	1
Raters' consistency	0.85						

Comparing and contrasting raters' comments: By comparing the raters' comments, and taking the number of vocabulary and discourse linkers, and complexity of sentence structures used by the participant's through the seven trials, it can be claimed that the participant improved in his speaking skill in terms of discourse linkers, lexis, and organization. The tables below indicate the participant's progress through the seven speaking trials. In order to show the participant's progress more clearly, the first and the seventh trials are presented here to make a direct comparison (Table 2). It should be mentioned that the first trial was regarded as the baseline data for comparison of subsequently acquired data.

Table 2. The raters' comments on the first trial

First rater 2/4	Second rater 2/4
• Basic Grammar (limited control on Grammar) • Using Chunks • Simple vocabulary • Limited elaboration	• Basic structure • Simple lexis • Lack of linkers • It was intelligible • It was on/off the task

The raters' comments on the seventh trial

First rater 3/4	Second rater 3/4
• Speech is generally clear • Fairly effective use of grammar and vocabulary, and fairly coherent expression of relevant ideas • Relationships between ideas may at times not be immediately clear. • It is organized	• More organized • Using more linkers • Speech is intelligible • Good use of vocabulary

Across the seven trials, only on the third trial, the co-raters' score did not match. The second rater believed that the participant did not elaborate well in the second part of the speaking and the answer was rather short, however, the first rater still believed that it deserved ¾ since it was well-organized, and the participant had used wider vocabulary range. Table 3 shows this mismatch.

Table 3. The raters' comments on the third trial

First rater 3/4	Second rater 2/4
• Frequent pauses-	• cohesive devices+
• Vocabulary+	• lexis+
• Simple structure-only one complex	• simple structure-just one complex
• More linkers and cohesive devices+	• short answers
• Better organization+	• second part not well-elaborated
• Wider vocabulary range	

*Vocabulary range:*In the first trial, mainly basic words were used by the participant, and some of which were repeated. Some examples are as follows:

When I was at high school and primary school I was reading book and always I was not spending my time on technology internet and TV. But now because of technology I spend and I don't read book. And one effect.... I can reach more information about everything what I want.

I prefer to have different friends than me. About characteristic ... I don't know when I have different friends I can get some experience from them. While I am spending my time with similar friend always I get bored.

However, the range of vocabulary became wider as the participant got closer to final trial. Also. words were not frequently repeated as in the fifth trial and the seventh, the last, trial:

In my opinion most famous person is Atatürk. He influenced on our country's people for independence and our education and our economy. Because we were in the First World War. We did not have power to fight against anybody and we did not have any literacy people and he made a lot of institutions to give education.

I agree with second statement because every student deserve education ıt does not matter they are poor or they are best student, intelligent and if they do not get education their children will nothave knowledge if best student get best education it will be some kind of class society.

I agree that in just some way. First young generation they do not have enough communication with each other and their family ties are weak and second they do not have job and they should compete with each other when they do it they feel themselves selfish and it makes them alone It forces them selfish people.

I have taken an English course before in some way it help me so it make me think different way when İ talk about something and secondly it separates me from other people and also I can gain more resource about newspaper on internet and other hand I can watch movie without subtitle or dublaj and the main factor in English I can speak with native speaker.

Cohesive devices: Cohesive devices were more limited in the first and second trials, and the participant mainly used coordinating conjunctions most of which were repeated. Below is an example for the repetition of basic conjunctions:

I was reading book and always I was not spending my time on technology internet and TV. But now because of technology I spend and I don't read book. And one effect.... I can reach more information about everything what I want. That's why it is positive effects. But sometimes it affect me negatively

In my opinion, most famous person is Atatürk. He influence on our country's people for independence and our education and our economy. Because we were in the First World War. (The fourth trial)

Also they do not have time to read book. (The fourth trial)

I agree that in just some way. First young generation they do not have enough communication with each other and their family ties are weak and second they do not have job and they should compete with each other when they do it they feel themselves selfish and it makes them alone It forces them selfish people (The seventh trial)

In some way it help me so it make me think different way when I talk about something and secondly it separates me from other people and also I can gain more resource about newspaper on internet and other hand I can watch movie without subtitle (The seventh trial)

Structure: The analyses revealed no noticeable improvement regarding the structures the participant produced during the course of seven weeks. Most of the sentence structures were of simple and compound ones, and when the participant produced complex structures, they were mainly limited to adverbial clauses. In addition, at times, the participant repeated the same structure while trying to support his idea. The following samples are from different trials:

When I was at high school and primary school I was reading book and always I was not spending my time on technology internet and TV. (The first trial)

Sometimes I imagine if I had large money what would I do. (The second trial)

We must analyze technology and information because world is going to forward. (The third trial)

If they do not get education their children will not have knowledge if best student get best education it will be some kind of class society. (The fifth trial)

Sometimes when I work on hospital some worker comes to doctors and present their medicine. (The sixth trial)

When I talk about something and secondly it separates me from other people. (The seventh trial)

For example they should have general culture and make their students aware of what is going on in world and they should have different hobby they can spend with their students enjoy time. They should know and learn about literature and history they can advise their students.

The interview

In the last week of the study, the researcher conducted a semi-structured interview to find out more on what the participant thought of the study, and how his ideas and thoughts corresponded with his reflections throughout the study. There were four main questions in the interview revolving around four main topics. The questions were:

Question 1

First one is…..you know after each speaking task, you transcribed your speaking, and corrected your own errors. Did you find this self-correction useful or not?
Question 2
How would you compare this Independent speaking task regarding TOEFL-IBT with daily life speaking? Are they similar or different?
Question 3
You know after each trial, I told you about raters' feedback. Did you find the feedback useful?
Question 4
Finally, have you found yourself more aware of the areas you need to improve when it comes to TOEFL-IBT Independent speaking tasks?
Can you give me an example? Can you elaborate which areas do you need to improve?

The answers of the participant to the above questions were analyzed qualitatively. The analysis revealed four main themes based on the participant responses to the questions. These themes are explained below.

Discovery learning: According to constructivists, learners build up new information on the existing information. Through discovery learning, learners discover the relationships between old and new information and through frequency of use these relationships are strengthened. In order to discover a new language feature, learners need to notice the new information through awareness raising or even explicit instruction. When learner's focal attention is drawn to a new feature in language, learning takes place. In this study, guided learning was used to draw the participant's attention on some new linguistic features such as discourse linkers.

Yes I found it very useful because it's better than someone teach you because when you teach yourself, it remains constantly.

According to the participant's response, self-awareness and self-correction of the errors were found to be helping the participant to store the corrected ones in his long term memory.

Comparison and contrast:

I think it's different because in daily speaking sometimes I speak some foreigners and my friends and if I make correct it's ok because I can correct in same time but in TOEFL must control time and it's fifty seconds and ten seconds and I don't have time to correct in same time, but if I practice always, I'll be able to eeee not make mistakes.

Importance of raters' feedback:

After each trial I made a correction part. I noticed that something I did… I made mistake, I correct myself but when someone else correct, looks over and tell me what's eeee lack of your conversation, The test taker mentioned that when he was not aware of the areas he needed to improve when it came to speaking and the raters' feedback was of a great assistance.

Metacognition:

I found out my eeeee what I eeee don't good at

I think its elaboration part and make more ,,,,complex sentence because I spoke sometimes with relative clause and the longer sentence and I think I must do that and may be use more different linkers because there is no just like also and that kind of there are more different linker word, I need to use them.

Discussion of the data

Although this study focused on all categories of scoring the candidate responses in the independent speaking part of the TOEFL-IBT speaking test, more emphasis was laid on the second and third scoring categories namely, structure, lexis as well as pragmatics. It should be noted that EFL teachers helping their students preparing for the TOEFL-IBT test sometimes give them templates to follow or sometimes record their students' speaking and then point out their mistakes, however, TOEFL-IBT test takers tend to repeat their mistakes corrected by their teachers. In this study, the researcher tried to make the test taker become more aware of his own speaking mistakes through the grades and feedback given to him by two raters based on the TOEFL-IBT speaking rubrics as well as his own reflection on his speaking. Carter (2003) defined language awareness as "the development in learners of an enhanced consciousness of and sensitivity to the forms and functions of language" (p. 64).

In other words, metacognitive knowledge and metacognitive strategies play a significant role in improving speaking skills. Flavell (1979) defines metacognition as someone's knowledge which concerns his/her cognitive processes and products. Horwitz (1988) believes that investigation of metacognitive knowledge can be useful, because it helps teachers become more aware of the different types of learners and better understated their needs and expectations, during the teaching process.

The present study aimed to make the participant become aware of both metacognitive strategies and metacognitive knowledge, and how they could both be used concurrently in EFL context. The test taker tried to find out his own errors through self-error correction. The learner then tried to self-manage his stress and anxiety as well as the time limit. In this stage, the learner became aware of the significance of time and stress management in the TOEFL-IBT independent speaking task. He also did his best not to repeat the mistakes he had made through self-correction and the feedback he received. The test taker also paid more attention to the language he produced in details. He tried to be independent of his/her teachers and become more autonomous. To illustrate, the participant mentioned this improvement in his final reflection:

Now I am able to make the test much easier. In the last trial, I wrote down some linkers and while I was speaking, I could use them. Also, I am able to elaborate my speaking. My speaking is not off the topic anymore.

It should be noted that the researcher only guided the test taker to identify the fossilized errors and the ones he was not aware of. According to Selinker (1972) , an English psychological linguist, when second language learners tend to retain items, rules and subsystems in their interlanguage while they are in the process of learning a particular target language fossilization occurs. These items, and rules require considerable effort to be eliminated regardless of explanation or instruction. Furthermore, the test taker learnt how to take responsibility when it came to decision making. In fact, he was the only person to make his own choices when it came to answering a question rather than using prefabricated chunks and speaking templates, which are generally taught and given to the candidates by teachers in TOEFL preparation classes. In a nutshell, this study was based on the cognitive approach rather than a mechanical and behavioristic one. Halfway through the study, it was noticeable that the participant used more linkers and discourse markers. He had also noticed the importance of discourse linkers in speaking by mentioning several times in his reflections that he needed to use more linkers to better illustrate his reasons.

This study aimed to evaluate the extent to which transcription of speaking tasks, would affect the improvement and development of TOEFL-IBT test takers' speaking skills in terms of accuracy and fluency. In preparation for the independent speaking tasks, the study took into account the learners' self - correction as well as teachers' guidance, and self-awareness of errors made by the student based on raters' feedback. Nobuyoshi and Ellis (1993) stated that the accuracy gained through self-correction and self-awareness tended to remain over a longer period. Stones (2012) did a similar study to the present study on IELTS exam test takers. The results revealed that the participants used their own corrections more than corrections done by their teacher; however, this study aimed to focus on independent speaking tasks of TOEFL-IBT, which are completely different from IELTS speaking parts. In IELTS speaking task, the examinees speak to the examiners, however, in TOEFL-IBT speaking tasks, the candidates do not speak to a real person; in contrast, in the first two independent speaking tasks, the candidates read and hear two questions on the computer, and they should answer and support their ideas by giving reasons. According to Thornbury and Slade (2006), speaking is a multi-sensory activity because speaking involves paralinguistic features such as facial expressions, body language, eye contact and pitch variation and so on and so forth, however, some of these features are absent when it comes to TOEFL-IBT independent speaking task. Another difference is that in the IELTS speaking section, test takers are not required to finish their speaking within a very specific time frame (only in section two, the examinees should speak between one to two minutes). The examiner decides when to interrupt the examinee; however, as mentioned above, in the TOEFL-IBT independent speaking parts, candidates have only 15 seconds to prepare their answers and 45 seconds to respond.

Conclusion

The researcher of this study investigated whether self-correction and self-awareness could help the participant of the study improve his speaking skills for the TOEFL-IBT speaking section. The results of this study provided evidence that through raising awareness and self-correction the participant's speaking skills improved in terms of vocabulary range, elaboration, using discourse linkers based on the raters' judgments, however, there was no noticeable improvement regarding sentence structure. Taking the raters' comments and scores into account, this study was found to have a positive effect on the participant's speaking skills. Based on the results obtained from the interview, it became clear that the participant became more aware of the areas he needed to improve for the independent speaking tasks. Moreover, the participant found self-correction an effective way to remember and avoid future errors later in his speaking. He also found the raters' feedback useful in that he became more aware of the ways the learner preferred to be instructed, which focused on discovery learning rather than rote.

Awareness raising and getting English language learners on the discovery-learning path can eventually smooth the path towards autonomous learning. Since learners' needs, motivation, and background knowledge are different, English language teachers and instructors, before planning their lessons, should take different ways and techniques leading language learners to discovery learning into account. One of the possible ways to achieve this goal is through reflection. Through reflection, language learners can develop and enhance their grammatical and strategic competence, which will eventually help them become more autonomous in their communicative competence. Another way to get language learners to become more aware of their errors and the areas they need to improve is providing them with feedback. Once, the learners have reflected on their speaking and

corrected themselves, it is time for English language teachers to provide them with feedback concerning the learners' correction and the points they have skipped correcting due to their insufficient English language knowledge or lack of enough attention. This, then, should be frequently done by the teachers in follow-up lessons to ensure the learners' improvement based on the given feedback.

Teaching, learning, and assessment are intertwined closely and should be seen as a continuous cycle. Teaching and assessment facilitate learning if teachers and learners make use of the feedback received from the assessment effectively that in turn may result in positive decisions and actions about learning. Since such decisions and actions are conscious ones, learners become more aware of what they are learning which in turn helps them feel more secure and less stressed particularly when it comes to speaking skills. Some teachers assess their learners to see how much they have improved at the end of a course. This kind of assessment focuses on the result; however, some other teachers tend to assess their learners' process of learning. Lantolf and Thorne (2006) called the latter one dynamic assessment. According to the Zone of Proximal Development proposed by Lev Vygotsky, there is a distance between what learners can do independently and without help, and what they can do with help and guidance from more capable learners or teachers. Ongoing and dynamic assessment along with feedback assist learners in moving from their actual developmental level to the level of potential development. Assessment and feedback can be viewed as scaffolding in this regard which helps learners achieve goals that would not be achieved without teachers' assistance. Vygotsky (1987) in "Thinking and Speech" laid emphasis on the importance of the ZPD in instructional practices. According to him, instructional practices should be in accordance with students' ZPD rather than the actual level of their development. In this case, learners feel far less stressed during the learning process and Dynamic Assessment and instruction are not independent from one another. Once learning has improved through ongoing assessment, learners start to notice the positive and beneficial effect of assessment in the learning process and their fears and lack of confidence in dealing with assessment medium start to fade away. Therefore, constant assessment in the long term would help learners view assessment as facilitating rather than a debilitating medium in learning.
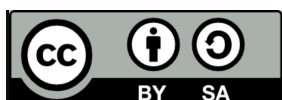
This study has some limitations that may lead to certain issues. Since this action study was limited to only one subject, the result of the study is not obviously generalizable. The other limitation of this study was that the co-raters were from two different countries namely the US and New Zealand in which English is spoken with different variations. This might have influenced the co-raters' judgments. Finally, the participant was not selected based on a proficiency placement test, but rather his English level that he had completed in the language school where this study took place.

References

Brown, A., Bransford, J.D., Ferraraand, R., &Campione, J.C. (1983). Learning, remembering and understanding. In J. H. Flavell and E. M. Markman (Eds.), *Carmichael s manual of child psychology Volume 1*, New York: Wiley.

Brown, H. D. (2001). *Teaching by Principles: An interactive approach to language pedagogy*. White Plains, NY: Longman.

Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics, 1*, 1–47.

Carter, R. (2003). Language awareness. *ELT Journal, 57*(1), 64-65. doi:10.1093/elt/57.1.64

Chaudron, C. (1977). A descriptive model of discourse in the corrective treatment of learners' errors, *Language Learning, 27*(1), 29-46.

Creswell, J. W. (2012). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research* (4th ed.). Boston, MA: Pearson.

Dormer, J. E. (2013). Improving speaking. accuracy through awareness. *Journal. of Adult Education,42*(1), 16-22.

Ekmekçi, E. (2016). Comparison of native and non-native English language teachers' evaluation of EFL learners' speaking skills: Conflicting or identical rating behavior? *English Language Teaching, 9*(5), 98-105. doi:10.5539/elt.v9n5p98

Ellis, G., & Sinclair, B. (1989). *Learning to learn English: A course in learner training*. Cambridge: Cambridge University Press.

Fukuda, Y. (2004). *Treatment of spoken errors in Japanese high school oral communication classes* (Unpublished Master's thesis). California State University, San Francisco, CA.

Farahani, A. A. &Salajegheh, S. (2015). Iranian EFL teachers' and learners' perspectives of oral error correction: Does the timeline of correction matter? *Latin American Journal of Content and Language Integrated Learning, 8*(2), 184-211. doi:10.5294/laclil.2015.8.2.6

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *American Psychologist, 34*, 232-911.

Gilman, R., & Moody, R. (1984). What practitioners say about listening: Research implications for the classroom.*Foreign Language Annals, 17*(4), 331-334. doi: 10.1111/j.1944-9720.1984.tb03236.x

Harris, D. P. (1974). *Testing English as a second language*. New York, NY: McGrawHill Book Company.

Hendrickson, J. (1978). Error correction in foreign language teaching: Recent theory, research, and practice. *Modern Language Journal, 62*(8), 387-398. doi:10.1111/j.1540-4781.1978.tb02409.x

Horwitz, E. K. (1988). The beliefs about language learning of beginning university foreign language students. *Modern Language Journal, 72*(3), 283–294. doi:10.1111/j.1540-4781.1988.tb04190.x

Lantolf, J.P., & Thorne, S. (2006). Sociocultural Theory and the Genesis of Second Language Development. Oxford: Oxford University Press.

Lemley, D. C. (2005). *Delayed versus immediate feedback in an independent study high school setting* (Unpublished doctoral dissertation). Brigham Young University, Provo, UT.

Long, M. (1977). Teacher feedback on learner error: Mapping cognitions. *On TESOL, 77*, 278-294.

MacIntyre, P., & Gardner, R. (1994). The subtle effects of induced anxiety on cognitive processing in the second language. *Language Learning, 44*, 283–305.

Madsen, H.S. 1983: *Techniques in testing*. Oxford: Oxford University Press.

Makino, T. Y. (1993). Learner self-correction in EFL written compositions. *ELT Journal, 47*(4), 337-341. https://doi.org/10.1093/elt/47.4.337

Mills, G. E. (2000). *Action research: A guide for the teacher researcher*. Princeton, NJ: Prentice-Hall.

Ming-chu, L., & Hung-chun, W. (2009). Perception differences of EFL teachers and students in grammar instruction and error correction. *English Teaching & Learning, 33*(1), 101-146.

Nobuyoshi, J., Ellis, R. (1993). Focused communication tasks and second language acquisition. *ELT Journal, 47*(3), 203-210. doi: 10.1093/elt/47.3.203

Rivers, W. M. (1984). *Communicating naturally in a second language: Theory and practice in language teaching.* Cambridge: Cambridge University Press.

Selinker, L. (1972). Interlanguage. *IRAL - International Review of Applied Linguistics in Language Teaching, 10*(1-4). doi: 10.1515/iral.1972.10.1-4.209

Stones, T. (2012). Transcription and the IELTS speaking test: facilitating development. *ELT Journal, 67*(1), 20-30. doi: 10.1093/elt/ccs051

Thornbury, S., & Slade, D. (2006). *Conversation: From description to pedagogy*. Cambridge, UK: Cambridge University Press. https://doi.org/10.1017/CBO9780511733123

Truscott, J. (1996). The case against grammar correction in L2 writing classes. *Language Learning, 46*(2), 327-69. https://doi.org/10.1111/j.1467-1770.1996.tb01238.x

Vandergrift, L. (1999). Facilitating second language listening comprehension: Acquiring successful strategies. *ELT Journal, 53*(3). doi:10.1093/elt/53.3.168

Vygotsky, L.S. (1987). Thinking and Speech. In The Collected Works of L. S.

Wenden, A. L. (2001). Meta-cognitive knowledge in SLA: The neglected variable. In M. P Breen (Ed.), *Learner contributions to language learning: New directions in research* (pp. 45). Harlow, Essex, United Kingdom: Pearson.

Zhang, Y., & Elder, C. (2014). Investigating native and non-native English-speaking teacher raters' judgements of oral proficiency in the College English Test-Spoken English Test (CET-SET). *Assessment in Education: Principles, Policy & Practice, 21*(3), 306-325. doi:10.1080/0969594X.2013.845547