

Peer Feedback: Is It Worth Listening to?

Yukie Saito

ABSTRACT

In peer feedback, students use criteria and apply standards to the work of their peers. As a way of enhancing learner autonomy, its benefits have been recognized from theoretical and pedagogical perspectives. Unlike studies on peer assessment in writing classes, it was in recent decades that researchers started paying attention to the value of peer feedback in speaking classes (De Grez, Balcke, & Roozen, 2012; Patri, 2002; H. Saito, 2008). These studies measured the reliability of peer feedback statistically but did not investigate it qualitatively. This study, which conducted both quantitative and qualitative analyses of peer feedback, finds that peer feedback achieved a certain degree of accuracy measured against teacher feedback but that there existed individual differences in the *specificity* and *descriptiveness* of their feedback.

INTRODUCTION

Although some empirical studies have reported the value of peer feedback in speaking classes (De Grez et al., 2012; Patri, 2002; H. Saito, 2008), the benefits it may bring to the classroom are not guaranteed unless students are capable of evaluating peers objectively and providing feedback of good quality (H. Saito, 2008). In my previous study (Y. Saito, 2012) that focused on students' attitudes toward peer feedback, a few participants showed dissatisfaction with their peers' perfunctory feedback and therefore, preferred teacher-centered feedback which they thought more reliable. Although the findings revealed that students' perceptions were to some extent influenced by feedback they received, the study did not fully address the issue regarding the quality of feedback.

One way of analyzing the quality of feedback quantitatively is calculating an agreement rate between peer assessment and teacher assessment as seen in Patri (2002), Saito (2008), and De Grez et al. (2012). When interpreting the results of these studies, it is crucial to consider whether or not training was conducted, and if so, what kind of training participants received. Patri (2002) had students experience a two-hour training session to familiarize themselves with assessment criteria for oral presentations and found that their evaluation about peers' presentation skills was comparable to that of the teacher. In De Grez et al. (2012), peer assessors were given formal instructions on oral presentation skills and the use of evaluation rubric. Despite a positive relationship between teacher and student assessment, they reported that peers and teachers interpreted the criteria in a different way. Somewhat different results were shown in H. Saito (2008), where control group was compared to treatment group that received additional 40-minute training. Although there were no significant differences between the two groups, training seemed to help students provide more relevant comments; students who received the training referred to more skills of peer performance.

As seen in the inconsistent findings above, the question regarding the reliability of peer feedback has not been sufficiently answered. In addition, none of the studies above, unfortunately, analyzed student feedback qualitatively. To address the gap in literature, this study not only examines the reliability of peer feedback statistically but also attempts to enrich the data by looking at several examples of actual student feedback. Now the question to be asked is what consists of "good" feedback and this will be defined in the following section.

Although teachers may instinctively know that vague and judgmental feedback such as "You did a good job!" is of limited value in promoting learning, the definition of effective feedback should be clarified. Black and William (1998, cited in Chappuis & Stiggins, 2002)

emphasize the importance of feedback being *specific* and *descriptive*. *Specific* feedback should pinpoint learners' strengths and weaknesses. In a discussion class, this can include examples of ideas shared in discussions and the target function phrases that the students use or do not use. Examples, which make feedback contextualized, are of great importance because "[f]eedback has no effect in vacuum: to be powerful in its effect, there must be a learning context" (Hattie & Timperley, 2007, p.82). *Descriptive* feedback should not only describe why certain performance is good or bad, but also guide students to better performance instead of just pointing out what they did wrong.

Based on these theoretical frameworks, the current study aims at answering two research questions:

1. How statistically different is student rating from teacher rating?
2. Can student feedback be specific and descriptive?

METHOD

Participants of this study were 17 male and 27 female first-year students enrolled in six of my English discussion classes in the spring semester 2012. The selection of the classes was made based on the level (level 2 and 3 classes were selected to minimize differences in proficiency among participants) and the class size (classes of nine students were excluded due to time constraints). It was assumed the majority of the students were accustomed to teacher-centered instructions in their previous learning contexts, and therefore, the focus of the first five weeks was for students to learn the basic skills of discussion and feel comfortable in a student-centered classroom. It was only in week 6 that peer feedback was introduced and this activity was conducted four times in total (in weeks 6, 8, 11, and 12). Data for this study were collected from check sheets used in week 12 and recordings of students' interaction during feedback sessions.

Students in pairs or groups of three were asked to monitor their peers discussing for ten minutes, tick boxes on a check sheet when they heard their partners using target phrases, and afterwards give feedback on the ideas discussed and use of functions. Brief instructions regarding how to give useful advice were provided in each treatment. For example, students were encouraged to include examples of functions phrases in feedback or to note down some key words of interesting ideas discussed. Though students were given one or two minutes to look through the assessment criteria, no specific training was conducted.

RESULTS

A handout used in week 12 was a yes/no check sheet for nine items: six functions and three communication skills. The average agreement rate for each student between their peer and teacher's rating was calculated by dividing the number of peer ticks by that of teacher's, and was .83. A statistical analysis was also conducted to see the internal consistency estimates of reliability. The mean of student rating was 5.11, indicating that a student, on average, used about five items out of nine target items. This number was slightly lower than the mean of teacher rating, 5.75. The value for coefficient alpha was .86, indicating satisfactory reliability.

In order to answer research question 2, student-student interaction during feedback sessions in the classroom was audio-taped by putting two IC recorders on tables. Although I have to admit that there were some constraints on capturing all the simultaneous interaction clearly, several examples from the recorded data can provide insight into actual student feedback. The extracts below, which are all verbatim, illustrate different degrees of *specificity*. Phrases in bold were already written on the sheet to help students construct their feedback. Examples of ideas and function phrases are underlined for this analysis.

- (A) **I enjoyed your discussion because** I can hear about some part-time job story. For example, your tutor work is, tutor work's income is so good. **You did a good job because** you used follow-up questions and said "Can I say something?" so it's good thing. (Level 2 female)
- (B) **I enjoyed your discussion because** your ideas were clear so I hear, I understand. **You did a good job because** you asked a follow-up question. For example, you asked Kaori, "Why did you choose this?" (Level 3 male)
- (C) **I enjoyed your discussion because** good discussion. For example, you used "for example." "Can I start?" "It's mainly because..." and so on. **You did a good job** of these three functions. (Level 3 male)

Extract A is the most specific; a literature-major female student gave feedback on the content by giving an example of her peer's part-time job as a tutor and on the use of functions by specifying the phrase her partner used. Although some students, as represented as in Extract A, were able to give feedback whose quality was almost equivalent to that of teacher feedback, there were a few students who found it challenging to give examples as indicated in Extract C. A male student struggled to complete the first sentence probably because he did not remember ideas discussed or even so, could not articulate them sufficiently. Extract B lies in somewhere between A and C. In this case, a male student was able to pinpoint his partner's strength by referring to a follow-up question she asked but failed to fully develop his reason why he enjoyed the discussion due to a lack of concrete feedback on the ideas discussed.

Next, shifting the focus to *descriptiveness*, this section will look more closely at how practical and useful student feedback can be, in particular, paying attention to the last part of feedback where students were trying to give advice to their peers. The most common type of advice, as shown in Extract D, was to name functions that peers did not use sufficiently. The phrase **"In the next discussion, try to..."** in bold was also provided on the activity sheet. Extract E is one of a few impressive examples. In this interaction, a female student was trying to give feedback on two things: her partner forgot to check understanding and used a disagreeing phrase. Not only did she point them out but advised when he could have used the phrases in the discussion, using his example to contextualize her feedback.

- (D) **Try to** connecting ideas and sharing experiences. (Level 3 female)
- (E) **In the next discussion, try to** more checking understanding after finishing your opinion and in addition, "Do you understand?" or "Do you follow me?" more. And you try to agree or disagree. For example, "I see your point, but I think succeeding in the society is more important..." (Level 3 female)

There is another interesting point worth mentioning in this analysis. Student feedback was not limited to the criteria set by the teacher but extended to a small voice ("Next time, big voice!"), repeated use of one function phrase ("You said 'I think' 'I think' 'I think' but you should say 'personally speaking, I think...' or 'in my opinion, ...'"), and politeness ("You used 'I disagree' but this is a little strong so please use 'I see your point, but...'"). In other words, students were able to make their feedback less-structured by being creative in giving advice to peers.

DISCUSSION

With the feedback criteria clearly set, the participants, in general, were able to quantify the number of target phrases accurately. By week 12, majority of the students seemed to learn the phrases by heart and feel comfortable using meta-language to give and receive feedback. Although specific training was not provided in this study, a relatively high agreement rate between the teacher and students was found and this is consistent with the findings of H. Saito (2008) that instructions on aspects of oral presentation skills were sufficient enough to achieve a certain level of correlation between the instructors and students and that training did not lead to more advantages. Since the criteria used in this study were easy-to-adopt – students were simply asked to do yes/no checking, having students repeatedly engaged in this activity might have worked as effectively as the instructions in H. Saito's study (2008).

The qualitative analysis of student feedback showed a somewhat different picture. All the students were instructed to give feedback on three points: ideas, things peers did well, and things peers should do more, and to give as many examples as possible to explain clearly. This instruction, however, did not guarantee equally valuable feedback from every participant. Some students were able to provide feedback that was as *specific* and *descriptive* as teacher feedback could be; others struggled to remember or verbalize examples of ideas and/or function phrases mentioned in discussion.

In order to answer how much reliability is required to make peer feedback activity meaningful, it is worth referring to different views on peer feedback. Some researchers argue that the benefits of peer assessment outweigh a certain level of discrepancy between student and teacher marks (e.g. Langan et al., 2005, cited in De Grez et al., 2012). On the other hand, those who emphasize the importance of the accuracy of feedback tend to see peer feedback as an outcome, and hence, try to train students to be reliable assessors. I agree with De Grez et al., who claim that a combination of both views is necessary, explaining “on the one hand we do not want students to take wrong actions based on low-quality feedback but on the other we want students to do something with the feedback” (2012, p.139). My interpretation of the findings of this study is although the degree of *specificity* and *descriptiveness* of peer feedback varied individually, the differences in its quality were not considerable enough to offset the value of peer feedback as seen in the high correlation between peer rating and teacher rating.

CONCLUSION

The results of this study reveal that despite individual differences in its quality among students, peer feedback achieved a high reliability compared to teacher feedback. Therefore, as a complement to conventional teacher-centered feedback, it can provide learners with meaningful learning opportunities.

REFERENCES

- Chappuis, S., & Stiggins, R. (2002). Classroom Assessment for Learning. *Educational Leadership*, 60 (1), 40-43.
- De Grez, L., Valcke, M., & Roozen, I. (2012). How effective are self- and peer assessment of oral presentation skills compared with teachers' assessment? *Active Learning in High Education*, 13(2), 129-142.
- Hattie, J., & Timperley, H. (2007). The Power of Feedback. *Review of Educational Research*, 77(1), 81-112.
- Patri, M. (2002). The influence of peer feedback on self- and peer-assessment of oral skills. *Language Testing*, 19(2), 109-131.

- Saito, H. (2008). EFL classroom peer assessment: Training effects on rating and commenting. *Language Testing*, 25(4), 553-581.
- Saito, Y. (2012). Peer feedback: Do our students react to it as positively as teachers wish? *New Directions in Teaching and Learning English Discussion*, 1, 4-83-4.88.