The Efficacy of Student Self-Assessments for Academic Speeches

Paul Nehls

Yokohama City University

Through primary research this paper provides insight into whether university-level student self-assessments of L2 speech performances can be an effective tool for helping students improve their presentation skills. The research focused on two possible self-assessment skills: student self-directed learning and personal goal setting. Self-assessment papers were collected from randomly selected student participants and the comments and suggestions they made on their assessments were compared to those of a group of independent speech instructors. The data was analyzed in regards to whether students' assessments agreed with instructor assessments and whether or not students assessed items that were not mentioned by instructor's assessments. Results of these comparisons are discussed and conclusions are made as to the effectiveness of student self-assessment for academic speeches in an L2 university classroom.

　本論は，大学生の第二言語によるスピーチの自己評価が，プレゼンテーション能力を上達させるための効果的な手段になるかどうかを考察する。この研究では、学習者の自発的学習と個人的目標の設定という二つの自己評価能力に焦点を当てた。無作為に抽出された学生の自己評価用紙を用い、自らの批評と今後への示唆を複数のスピーチ担当の教師による評価と比較した。データは学生の自己評価が教師の評価と一致しているか、教師が含めなかった評価項目を学生が自己評価しているかを元に分析し、その結果を議論する。最後に、大学の第二言語教育現場での学習者によるアカデミック・スピーチの自己評価の有効性について結論をまとめる。

Self-assessment has been a tool used by public speakers and educators for decades, and since the 1980s self-assessment as a tool for L2 learners has received attention from researchers. It has been established that self-assessment can be used successfully in L2 classrooms (Brantmeier, 2006; Little, 2005; Rivers, 2001) and that it is an effective way to facilitate student learning (Ekbatani, 2000; Nunan, 1988). Furthermore, as summarized by Todd (2002), self-assessment has been shown to be a prerequisite towards developing self-directed learning and can increase motivation and goal orientation in learning. Todd further concluded that self-assessment can look into questions of learner effort and belief, which are areas of learning assessment language teachers cannot make for their students.

The majority of studies into L2 self-assessment have looked at student competences in regards to reading, listening, and grammar as determined by written placement tests, with some using further posttests (Brantmeier, 2006; Hargan, 1994; LeBlanc & Painchaud, 1985; Ross, 1998). Less common are studies that look at speaking competence as it is a more difficult skill to test for a variety of reasons, such as the difficulty in creating a clear definition of *competence* or *fluency*. As Butler, Eignor, Jones, McNamara, and Suomi (2000) asserted, "The available research does not provide a firm foundation for constructing a specific test of speaking as part of second language academic communicative competence." The current state of the research is such that at the completion of their literature review, Babaii, Taghaddomi, and Pashmforoosh (2015) stated, "Learner's criteria when assessing their own speaking ability as examined against teachers' rating criteria for L2 speaking have been rather underexplored" (p. 4).

Self-assessment has a long tradition of use and value in L1 public speaking contexts such as business and sociology and there is research to bear this out (Iwamoto, 2007). There are also a few studies that have focused on L2 self-assessment of speaking competence. Babaii, Taghaddomi, and Pashmforoosh (2015) used monologues on general topics as part of an IELTS (International English Language Testing System) style test and concluded that when provided with clear assessment criteria that were discussed by teachers and students, the learner's self-assessment was more likely to align with the teacher's assessment. Chen (2008) also found that through collaborative assessment students' self-assessments began to align more with the teacher's as they progressed through the speaking activities. Although these studies and others like them have looked at self-as-

sessment of L2 student performance during speaking tasks, they have not investigated self-assessment of more formally prepared (and presumably practiced) speeches or presentations by L2 students of English.

This paper is a report on a case study meant to address this gap in L2 self-assessment research. A close comparison was made of student self-assessments and a set of instructor assessments to try to determine the effectiveness of student self-assessments within the context of L2 prepared speeches and presentations.

Methodology
Research Questions

Within the context of L2 speaking competence, several questions will be explored through a comparison of written assessments done by students and instructors:

RQ1. When assessment is not determined by tests but rather by instructor comments and suggestions, how do student self-assessments and instructor assessments compare?

RQ2. Do student perceptions of progress match the instructor's assessment of progress?

RQ3. If self-assessment is meant to achieve self-directed learning and greater learning awareness leading to personal goal setting, do students note and assess skills and behaviors that their instructors do not?

Answers to these questions will be used to discuss if student self-assessment is a reliable or useful tool in the context of advanced L2 speaking and presentation exercises.

Participants

The participants in this study were part of an advanced English course at a Japanese public university. The focus of the course is *academic speaking* with the goal being to prepare the students to take content courses taught in English at the university, as well as be able to undertake courses abroad in English-speaking countries. To take part, students must have passed required prerequisite English classes and have test scores of either 520 or more on TOEFL ITP or an equivalent score on the TOEFL IBT, IELTS, or TOEIC examination. The majority of students enter the class with the requisite scores on the TOEFL ITP or TOEIC tests. However, because these examinations do not test communicative competence, the speaking ability and experience of the students in the course can vary greatly.

To represent the instructor assessment for this study, three instructors from the same university were selected to participate. Two criteria were used in the selection. First, the instructors had not taught any of the student participants. This was done to limit bias towards any of the students based on a prior relationship, and because they had not previously evaluated these students in any capacity they could watch the video recordings without bias based on past performance. The second criterion was understanding of and experience in using the curriculum. All instructors selected for the study had taught or were currently teaching the same academic speaking course using the same curriculum; they were all experienced in using the established criteria of assessment of student speeches and presentations. Therefore, although each instructor demonstrated a slightly different style of commentary, all of the instructors were looking for and commenting on the same issues.

Procedure

Throughout the semester in this course, students give several prepared speeches. After each speech, before receiving any feedback from the instructor, students watch a video recording of their performance and fill out a self-assessment form (see Appendix). Only after the students have submitted their self-assessments do they receive the instructor's comments. During the course it is repeatedly stressed to the students that this self-assessment requirement is a core learning component of the class and should be taken as seriously as possible. As asserted by Aryadoust (2012), "The more the students learn about the value of self-awareness and the more emphasis teachers attach to it, the more precise the self-assessment becomes during the course" (p. 3). Self-assessment questions are open-ended to encourage students to notice as much as possible with little predetermined direction or expectations. The focus of the course is on the presentation (public speaking) skills. Through lectures and classroom activities students practice and analyze their general public speaking skills such as eye contact, body language, vocal projection, and general fluency. Little attention is paid in the class to grammar or pronunciation, so these comments rarely appear on the student's self-assessments.

Course participants were notified that they could volunteer their self-assessments and video recordings to be used for research purposes and consent was given by 17 students from four different classes. From these participants, 10 students were randomly selected.

Videos of these students' first three speeches performed in the class and their subsequent self-assessments were used. The number of participants was later reduced to nine due to an unfortunate file corruption of one of the participant's videos.

The videos were given to the three instructors selected to perform the assessments.

The instructors independently watched all of the videos (27 videos in all) and wrote assessments of the students' performances. Instructor assessments were focused on the class curriculum guidelines. The instructors looked for and commented on general performance criteria as well as perceived improvement (or lack thereof) by the participants through the progression of the videos. Instructors did not comment on grammar or pronunciation.

Comments made on the student self-assessments and the instructor assessments of the second and third videos were directly compared and analyzed for similarities and differences. The first video of each student was used to establish a basis from which to determine improvement or lack thereof. The final analysis is based upon a comparison of 18 self- and instructor assessments of the 18 video recordings.

Results

Agreement/Disagreement Between Instructors and Students

When comparing the instructor and student assessments, two items were found that related to the research questions. In all the assessments, instructors and students noticed and commented on the same issues (e.g., eye contact and vocal projection) a total of 23 times. In 15 of the instances, the instructors and students agreed that the issue was either essentially positive or essentially negative. Examples of positive comments included some skill that had been done well (such as eye contact) or proved effective with the audience (such as asking rhetorical questions) or something the student had done that showed improvement. Examples of negative comments included some skill that had not been done well, needed improvement, or proved ineffective. There were also eight instances when instructors and students noticed the same issue but disagreed on whether it was positive or negative. For example, one student talked specifically about the clarity of their[1] message in their second speech saying, "I was able to speak more understandable and descriptive than previous" and "I could spend more preparation for this speech and make the speech clearer." The instructor comments regarding clarity noted that the speech was largely difficult to understand: "He was difficult to follow," "Intelligibility was lacking," and the speech apparently "had no clear conclusion." Comments were determined to be in agreement or disagreement only when the language was sufficiently similar. For example, if both commented on "eye contact," these comments were used to determine agreement or disagreement. However, if the student referred to themselves as their body being "towards the audience" and the instructor commented on "eye contact," this was considered as two separate issues. The instances of agreement and disagreement are shown in Table 1.

Table 1. Student Comments That Agreed or Disagreed With Instructor Comments

Student	Agreement	Disagreement
A	3	1
B	0	3
C	3	1
D	3	1
E	2	0
F	1	0
G	0	1
H	2	0
I	1	1

Following are examples of student comments that agreed with instructor comments. One student commented on their second speech, "I tend to repeat the same action when I tried to think what I wanted to say." This repetitive body language was also noted by the instructors in their assessments. The same student commented on speech number three regarding their lack of eye contact, "I had a lot of looks at the paper . . . my eye contact didn't look normal . . . . The topic was a bit difficult for me . . . . It was difficult for me to talk to the audience." The instructors also commented to the effect that the student's eye contact was insufficient and had not improved from the previous speeches. Another student made a comment that matched the assessments by all three instructors: "This outline method helped me to talk by looking at the audience and with my own language . . . . It was better than the last time because I was not holding the paper." One of the instructors noted, "This speech was better organized and eye contact improved—less reading from a paper."

There were eight instances when instructors and students noticed the same thing yet disagreed about whether it was positive or negative or whether it showed improvement or lack thereof. In one case, a student noted on the assessment of their third speech, "I could look at the audience longer time than the first 2 speeches." However, all three instructors' assessments were that the student's eye contact was *not* as frequent as in the previous speeches and that the student had in fact regressed in terms of using good eye contact.

contact. About their third speech, another student commented, "I was able to speech more understandable and descriptive than previous." However, the three instructors were unanimous in their assessment of this speech that it was the most lacking in coherent structure and vocal variability (tone and cadence). The instructors remarked that the student failed to provide emphasis to indicate what parts were important or marked transitions, thereby making the speech difficult to follow. One instructor noted: "What happened? This was the hardest speech to understand and seems to be the most lacking in proper preparation." There were some instances when students wrote negative assessments and instructors wrote positive assessments. One student wrote about their second speech, "I didn't talk well. I need more practice!" The instructors, however, reacted on the whole positively. One noted, "Better introduction, more animated with gestures and more friendly (than the first speech)."

Student Noticing

The assessments were also analyzed to determine how often students and instructors found different behaviors or skills to comment upon. This is relevant in determining whether or not the self-assessments are encouraging student self-directed learning.

Overall it was found that students noticed and commented on behaviors and skills that did not receive instructor comments a total of 25 times. These are shown in Table 2.

Table 2. Items Students Noted That Were Not Mentioned by Instructors

Student	Number of times
A	0
B	2
C	2
D	4
E	4
F	3
G	5
H	3
I	1

The following are items from student self-assessments that were not noted by the instructors. The four things mentioned by Student D included both negative and positive items. For the same speech they stated positively, "My speed was more natural," and negatively, "My body language was random." Student F noted two negative behaviors: "random hand movements" and "not enough eye contact." Most of the comments made by the students were criticisms of their performance, indicating skills the students wished to improve. However, not all of the comments were negative. This suggests students are open to noticing encouraging things, not only critiques.

Instructors commented 9 times about things that were not noted by the students. These were also a mix of both negative and positive observations such as "speaking rate was robotic" and "appeared very relaxed and confident."

Discussion

There are two observations that can be made from this case study. First, although in the majority of instances instructors and students agree in their assessments of speech performances, they do not always agree. This disagreement is informative and can be seen as an opportunity to open a dialogue between instructor and student to further analyze and assess ability. The fact that this disagreement happens is something that instructors should be aware of so they can address the issue appropriately. In cases in which the student self-assessment is positive and the instructor's is negative, the instructor's comments could be discouraging to students. Reasons for this disagreement should be explored. When students' comments are negative and those of the instructor are positive, this is an opportunity for the instructor to potentially boost a student's confidence.

The fact there was disagreement in assessments is also important for another reason. It suggests that the students were taking the opportunity to evaluate themselves honestly and were not feeling pressured to fill out their self-assessment form as they imagined their instructor might wish them to answer. In terms of self-directed learning and the setting of personal goals (see Todd, 2002), it is an encouraging sign that indicates self-assessments of speech abilities are indeed demonstrating usefulness for the students.

Furthermore, in answer to the research question of whether or not self-assessment is achieving self-directed learning and greater learning awareness, this study offered a clear "yes" for an additional reason. In 18 assessments, the students noted 25 instances of skills and behaviors that the instructors did not comment on. One interpretation is that perhaps the students were indeed cultivating a personal-assessment awareness beyond what their instructors were providing, and this could in turn prove to be instrumental in achieving more self-direction and goal setting in their learning.

Although comparing the student and instructor assessments was time consuming, it was not particularly arduous work. The possibility certainly does exist for teachers in L2 contexts to conduct similar comparisons of their own assessments with those of their students. Instructors need not bring in independent instructors as was done in this study. If an instructor has completed their own analysis before the self-assessments are collected, the instructor could make a comparison before giving the student their comments. This is a simple means for instructors to compare whether or not they and their student agree, whether or not they are noting the same issues, and how the dual assessments can be used to better address a student's progress, thus increasing learning potential. This process opens up new avenues of conversation and learning between the instructor and student in classes where speaking is the focus.

As noted previously, given the apparent lack of published work on L2 self-assessment for prepared speeches and presentations, further research should be conducted in this area. This study was limited in scope in that as a case study no attempts were made at statistical analyses to compare student and instructor assessments. The small scale of this study also means a statistical analysis would not have been accurate. Furthermore, the criteria and questions used for assessment were relatively open. In the future, a larger scale study that would allow for statistical analysis, using a methodology focused more closely on particular issues, skills, or techniques, is necessary.

Conclusion

In this study I compared and analyzed student assessments and independent instructor assessments. The evidence indicates that self-assessments done in an L2 academic speaking course are indeed effective in encouraging self-directed learning, and perhaps also enhanced personal goal setting by the students. The evidence also suggests that instructors of speaking courses can improve assessment dialogues with their students by taking the time to compare their assessments with that of their students. It is recommended that student self-assessments and comparison of student and instructor assessments be used as a component of L2 academic speech and presentation courses to encourage more student-directed learning and goal setting, as well as better learning opportunities through improved student and instructor assessment dialogue.

Note

1. In this paper *they*, *their*, and *themself* are used as singular pronouns that are not gender specific.

Bio Data

Paul N. Nehls is an instructor at Yokohama City University. He holds an MA in applied English linguistics from the University of Wisconsin-Madison and has been accepted to a PhD program at Yokohama National University starting April 2016.

References

Aryadoust, V. (2012, May 29). Reliability of second language listening self-assessments: Implications for pedagogy. *English Language Teaching World Online (ELTWO), 5.* Retrieved from <https://blog.nus.edu.sg/eltwo/2012/05/29/reliability-of-second-language-listening-self-assessments-implications-for-pedagogy-2/>

Babaii, E, Taghaddomi, S. and Pashmforoosh, R. (2015). Speaking self-assessment: Mismatches between learners' and teachers' criteria. *Language Testing,* 1-27. Retrieved from <http://ltj.sagepub.com/content/early/2015/06/30/0265532215590847.abstract>

Brantmeier, C. (2006). Advanced L2 learners and reading placement: Self-assessment, CBT, and subsequent performance. *System, 34,* 15-35.

Butler, F., Eignor, D., Jones, S., McNamara, T., & Suomi, B. (2000). TOEFL 2000 speaking framework: A working paper (ETS Research Report No. RM-00-06). Princeton, NJ: Educational Testing Service.

Chen, Y. (2008). Learning to self-assess oral performance in English: A longitudinal case study. *Language Teaching Research, 12,* 235-262.

Ekbatani, G. (2000). Moving toward learner-directed assessment. In G. Ekbatani & H. Pierson (Eds.), *Learner-directed assessment in ESL* (pp. 1-11). Mahwah, NJ: Erlbaum.

Hargan, N. (1994). Learner autonomy by remote control. *System, 22,* 455-462.

Iwamoto, N. (2007). Is self-assessment possible? Japanese students self-assessment in their course work. *JAELL, 17,* 65-79. Retrieved from <www.jaell.org/gakkaishi17th/Noriko%20IWAMOTO.pdf>

LeBlanc, R., Painchaud G. (1985). Self-assessment as a second language placement instrument. *TESOL Quarterly, 19,* 673-687.

Little, D. (2005). The Common European Framework and the European language portfolio: Involving learners in their judgements in the assessment process. *Language Testing, 22,* 321-336.

Nunan, D. (1988). *The learner-centered curriculum.* Cambridge, UK: Cambridge University Press.

Rivers, W. P. (2001). Autonomy at all costs: An ethnography of metacognitive self-assessment and self-management among experienced language learners. *The Modern Language Journal, 85,* 279-290.

Ross, S. (1998). Self-assessment in second language testing: A meta-analysis and analysis of experimental factors. *Language Testing, 15,* 1-20.

Todd, R. W. (2002). Using self-assessment for evaluation. *English Teaching Forum, 40*(1), 16-19.

Appendix

Sample Self-Assessment Paper

Self Evaluation Video Recording #2   Name_____

- What positive things did you see or hear while watching your video?
- What negative things did you see or hear while watching your video?
- What did you see or hear that has improved since your previous speech(es)?
- What did you specifically notice about your body language during this speech?
- What would you like to focus on improving for your next speech?