

# Peer evaluation in a Filipino L2 speech class: "Speaking the unspeakable?"

Mildred A. Rojo-Laurilla  
*De La Salle University-Manila, Philippines*

---

## ABSTRACT

This is a preliminary study that looked at how Filipino students of English One at De La Salle University-Manila, Philippines have assessed the speech performances of their peers in the form of written feedback. The speech band was utilized in looking at the areas of speech where the students had to frame their evaluation within a separate narrative critique submitted to the teacher. The paper attempted to describe how the students viewed or evaluated their peers' performance by looking at how they wrote their comments and determined whether these peer narrative evaluations supported the numerical evaluations given for the peer.

Findings of the study indicated that students mostly evaluated through a comparison and contrast of their positive and negative impressions of the peers' performances. Positive comments for the peer generally focused on the following items: content, organization and vocal expression, while negative comments for the peer focused mainly on the following elements: categories of physical expression, vocal expression and others. The results also indicated that students provided a more positive perception of their peers' performances. The study was able to provide an insight into how the narrative peer evaluation may be used in a second language (L2) speech classroom. The written peer evaluation allowed the students to "speak" on matters that were otherwise "unspeakable" in class, even if it initially meant addressing the teacher, rather than the peers themselves.

---

## Introduction

This is a preliminary study that looked at how Filipino students of English One at De La Salle University-Manila, Philippines have assessed the speech performances of their peers in the form of written feedback. Although the study was designed in such a way that it would focus both on self- and peer assessment, what will be reported here focuses primarily on peer assessment.

English One is a Reading-into-Writing class that uses the process approach to expository and argumentative writing. English One (ENGLONE) has both the writing and speech components. In the speech component, students do not videotape their performances. They evaluate their peers' performances by accomplishing an evaluation sheet. Students mark the peers by providing them with a numerical feedback on their performances. This aspect of evaluation is similar to what is being done at the advanced oral communication classes (except that the videotaping and writing of a self-critique are additional and inherent components of the assessment procedures in the advanced class). This evaluation

sheet is based on a speech band that guides students in their peer evaluation process. A student typically delivers a speech worth 100 points. As a departmental policy, the 80% of the grade comes from the teacher while the remaining 20% comes from the peer. The peer grade is computed by adding all the raw scores from the evaluators divided by the number of evaluators to get the average score. The teacher then multiplies this by .2 to get the 20%. This, then, would be added to the teacher's grade that has already been transmuted to 80% (the raw score multiplied by .8).

As part of the students' training and on top of the speech band-based evaluation, teachers may ask the students to write a written evaluation of either their or their peers' performances. Teachers mark the students for submitting the paper, but the main purpose of having them write an evaluation is to provide feedback to the peers, or if a self-evaluation is solicited, this is to enable self-reflection on the learners' part. The other paper requirements for the speech component include outlines, exercises, and quizzes. There were several occasions at faculty meetings when teachers articulated observations that students relied more on the numerical grades from the peer, than from a paper-based narrative evaluation. If oral evaluation is received, it comes mostly from the teacher. If not from a one-on-one consultation with the teacher, or from the teacher's notes on the individual performances, evaluation may come from a general assessment of the over-all performances as discussed in the classroom, rather than through an in-depth and personalized peer assessment method. The study looks at this written peer feedback solicited by the teacher, as provided by the speech students to their peers.

### **Issues in peer evaluation**

As mentioned earlier, a common activity in a process-driven approach to composition writing is peer evaluation, particularly the peer-review technique or peer editing (Mangelsdorf, 1992). This allows for participation in the writing classroom, authentic feedback, making allowances for the opinion of others, and hopefully improvement in composition.

Unfortunately, it is not always the case especially in an L2 context. In Sengupta's (1998) experience, student writers may not always trust their peers, and that peer evaluation often turns out to be an exercise in futility because students are more preoccupied with figuring out easy ways to complete the evaluation sheets than evaluating the text (Sengupta, 1998, p. 21). The effect is that learners pay only lip-service to the task (Mangelsdorf, 1992, p. 280 as cited in Sengupta, 1998). Such a situation is not exclusive to the Chinese learners in Sengupta's study, but it has to be ascertained as far as DLSU's ENGLONE students are concerned.

But whether peer evaluation is within the context of a writing or a speech class, it has been found to be an important component of assessment for learning. It is, after all, the teachers' task to provide opportunities for pupils to develop the skills required for peer and self-assessment (Northern Ireland Curriculum, n.d.).

Patri (2002) observed that studies on the validity of self- and peer-assessments

have revealed some contradictory results. For example, Stefani (1994) and Williams (1992) found a close agreement between self-ratings and teacher ratings, and between self and peer assessments. Others, however, revealed a tendency for the learners to over- or underestimate their own or their peers' performances (e.g., Boud & Tyree, 1979, Heilenmann, 1990). This, in turn, may affect assessment (Patri, 2002; Tsui & Ng, 2000). This under- or over-marking may be due to the fact that it is difficult for learners to critique their friends and even themselves. It was also noted that other factors such as experience, maturity and simplicity of task also play an important role in the assessment process, thus affecting the reliability of self- and peer-assessment (e.g., Stephani, 1994; Williams, 1992). But other studies implied that self- and peer assessment is a valid and reliable measure of communicative language ability (Ross, 1998; Shore & Shore, 1992).

The ways by which peer observation is conducted may also play a role in the outcome of the assessment. Cosh (1999) identified different models of appraisal through peer evaluation and asserted that the existing models are detrimental to the observed's confidence. This is because rather than observing to learn from others, observation becomes an opportunity to assess and criticize the styles and methods of others. Although the context of her proposed model addresses the teacher more than the students in the evaluation process, the model highlights the need to encourage self-reflection and self-awareness—in short, the teacher's own development rather than the presumed ability to develop the teaching of one's peers or colleagues (Cosh, 1999, p. 25). This can be adapted in the speech class, where aside from providing oral feedback about how the peers performed, observation and peer assessment should also be a way for the students to discover their own strengths and weaknesses based on what they have seen from the others. The written peer feedback will hopefully spark the students' awareness of what aspects need improvement in their performances as they go through the process of articulating what they have observed from the peers.

In relation to the above issues, the study argues that written peer evaluation is also worth looking at in the speech class. Admittedly, the concept of peer evaluation in the writing class may be more adaptable especially when using the process approach because of the very nature of the process approach itself: peer and teacher feedback is an essential part of revision (Tsui & Ng, 2000). But how can it be usefully done in a speech class that is not necessarily using a "process approach" (except perhaps in the conceptualizing and drafting of outlines)? Although Filipino English One students are trained to "peer evaluate" numerically and non-numerically, the researcher argues that students must be aware of the value of assessment other than the numerical marks given by both peers and the teacher. Thus, it is of interest to explore written peer feedback to determine if it would also be beneficial as an assessment tool.

In the process approach to writing, feedback can be given on form and/or content (Ashwell, 2000), but this may not be completely applicable in the speech class as there may be a different set of criteria to be observed and considered. Hence the ways by which peer evaluation may be conducted must rely on a different rubric since characteristics of a speaking activity differs from a writing activity. The value of using the narrative-based peer feedback in a speech class

aside from the numerical marks has not been explored deeply in the DLSU-context, thus the impetus for the present exploratory study.

To answer these issues and concerns, the study explored the concept of the written evaluation with the general goal of documenting this experience of the Filipino L2 learners within a speech class setting. Specifically, it aims to find answers to the following questions:

1. How do students express their observations on paper? What linguistic markers do they use when articulating their narrative comments on their peers' performances?
2. Do their peer narrative evaluations support the numerical evaluations given to their peers?

## **Methods and procedures**

### *Participants and methodology*

The participants were English One Speech component students of the researcher in the 2nd Term of SY 2002-2003. Four sections of the Speech component class with about 40 students each were part of the study. The Sections were composed of: freshmen International Studies students (Section I); freshmen Accounting students (Section II); freshmen Computer Science students (Section III), and another section of freshmen Computer Science students (Section IV).

The students had completed their first speech requirement (two of the sections delivered descriptive speeches, while the other two sections delivered demonstration-of-a-process speeches) and were told to submit a reflective self-evaluation of their performances. Aside from evaluating themselves, they were also requested to evaluate their other classmates (the focus of the research). Five special evaluators per speaker were assigned randomly. They were told to use the speech band as a guide in developing their narrative comments. The speech band is a commonly recommended evaluation guide for public speaking classes that utilized descriptors for the criteria that include: Organization, Content, Physical Expression, Vocal Expression and Language/Description and rated on a five-point scale (poor, fair, average, good and excellent). They were told that their comments should focus on these major areas and they were also allowed to include other insights, if they had, on the said performances (see Appendix for the actual speech band used). However, the researcher did not recommend to the students a specific format (e.g., the students may write in paragraph or essay form, while another may write using the bulleted or numbered format) to allow flexibility in expressing themselves.

The students were asked to submit the evaluations a week after their speech or performance dates. This second narrative-based evaluation intended for the peer is the focus of this paper. The aspect of self-evaluation within the same research framework and methodology will be discussed in a separate report.

### *Data collection*

To achieve uniformity in the data analysis, the researcher chose ten speakers randomly from each of the four sections. Of the five special evaluators (SEs) for each selected speaker, the researcher chose the narrative peer evaluations of only two randomly selected SEs to be considered for the study. This is to make the data size more manageable. The data set generated a total of 80 peer evaluations.

### *Data analysis*

All evaluation reports were read several times by the researcher to determine if there were any recurring patterns in the students' narratives. Repeated readings revealed that the comments were either positive or negative towards the speaker's performance in terms of a given criteria, which became the major basis for creating categories of the students' comments.

Each evaluation sheet was consequently coded for positive or negative comments. In addition, comments that included both a positive and a negative remark were labeled separately under "backtracking or hedging" because it seems that a positive comment was initially given, but the student then negated it or backtracked by replacing it with a counter-comment. It may also be considered hedging because the student could not make up his or her mind about the performance and so the evaluator seemed ambivalent whether to consider such a performance as positive or not, hence, hedging the assertions.

Descriptive statistics (such as frequency) constitute the analysis (especially for Research Question 2), as well as illustrations on the types of comments, and the corresponding interpretation of the results.

## **Results**

The first research question is *How do students express their observations on paper?* A sub-question arising from this first question is *What linguistic markers do they use when articulating their comments on their peer's performances?* The students' narratives were observed to have a predominantly recurring pattern that their comments were either positive or negative toward the speaker's performance. In determining the linguistic markers, the researcher considered these positive or negative dimensions based on topic-theme of the comment and through some adjectival or adverbial descriptors. An example of a positive comment is "... the speaker did well...", while a negative comment could be in this form: "she forget<sup>1</sup> what she was supposed to say... she stuttered..."

Upon tallying the frequency of occurrences, the researcher then re-coded them according to the criteria set for the speech itself. Table 1 shows the relative frequencies of positive peer evaluations. A total of 233 positive comments were obtained from the data. Of this total, the majority (N=57) of the positive comments were under the "content" criteria, followed by "organization" (N=46) and "vocal expression" (N=37). More positive comments were obtained from Section III (N=71) followed by Section IV (N=65).

---

<sup>1</sup> All student comments cited in this paper are presented without any editing.

**Table 1**  
Positive comments for the peer

Category	Sections				TOTAL
	I	II	III	IV	
Content	18	13	13	13	57
Organization	7	11	15	13	46
Physical Expression	6	6	14	5	31
Vocal Expression	9	6	11	12	37
Language/Description	9	2	12	12	35
Others	5	6	6	10	27
<b>TOTAL</b>	<b>54</b>	<b>44</b>	<b>71</b>	<b>65</b>	<b>233</b>

The data seem to indicate that students focused on the content of the speech more than on the other aspects. Organization comes next. Students were able to detect whether the peers' speeches were proceeding according to what they had expected or not. The special evaluators were given copies of the speaker's outline to help them evaluate organization during the performance. Vocal expression is also another aspect of speech that the students seemed to focus on.

Some examples of positive comments from the peer evaluation are:

- **Content**  
*"Her topic is how to go back to camp when a camper got lost in the wilderness and how to use the handy compass she had. She also explained that the arrow of the compass always points to the north because of some theories involving the forces of nature. It was a good topic and she did very well."*—Giselle, Sect III  
*"In my opinion, the speaker did well in presenting her topic. She was very confident and her explanation of the process, she convinced us of what she was trying to prove. She taught us tips on how to play the instrument and she definitely made me feel that I would like to try it out for myself."*—Nikolai, Sect IV
- **Organization**  
*"Her speech was somewhat organized starting from the introduction down to the conclusion. The transitions were also somewhat clear."*—Love, Sect I
- **Vocal Expression**  
*"She exemplified effective gestures as well as appropriate facial expression and a varying voice tone."*—Rachel, Sect II

Now let's turn to the relative frequencies of the negative comments for the peer (see Table 2).

**Table 2**  
Negative comments for the peer

Category	Sections				TOTAL
	I	II	III	IV	
Content	8	9	2	6	25
Organization	5	8	2	4	19
Physical Expression	16	14	3	11	44
Vocal Expression	11	14	2	9	36
Language/Description	3	17	0	3	23
Others	7	10	6	7	30
<b>TOTAL</b>	<b>50</b>	<b>72</b>	<b>15</b>	<b>40</b>	<b>177</b>

A total of 177 negative comments for the peer were gathered from the data, with more comments given for physical expression (N=44), followed by vocal expression (N=36) and others (N=30). Majority of the negative comments were cited by students from Section II (N=72), subsequently followed by Section I (N=50).

It appears that errors or perceived characteristics that impede successful delivery still fall under what the students actually see. These are physical mannerisms or movements that accompany speech.

Some examples of negative peer evaluation comments are:

- Physical Expression
 

*"Her facial expressions are limited but it somehow replaces her lack of voice projection."—Teresa, Sect II*

*"It seemed that he was guessing and unsure of what he's explaining. This is visible in his actions. He kept on scratching his chin, for me, sign of guessing. His expression was blank."—Janika, Sect IV*
- On Vocal Expression
 

*"She 'reads' too fast that in some point her tone becomes monotonous."—Teresa, Sect II*

*"The speech could have been delivered perfectly if she just spoke loud enough and delivered it with emotions."—Carl, Sect II*
- Others
 

*"Her hair is partially covering half of her face which is rather distracting for her audience. She did not share much of her own insights about her topic. She merely based everything on maybe what she had read or researched."—Christian, Sect II*

Figure 1 summarizes the relative proportions of the students' positive and negative comments.

Over all, students rated their peers more positively than negatively based on the frequencies obtained (i.e., 233 positive comments versus 177 negative ones). Meanwhile, there were instances of backtracking or hedging in the comments. These are presented in the table below per section (Table 3). The data found 45 instances of backtracking and hedging. Section IV (N=29) listed the most number of hedges among all the sections.

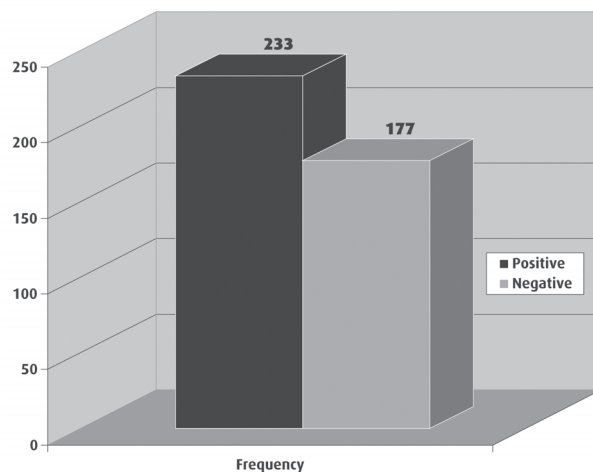
Here are some examples of the backtracking or hedging highlighted from the evaluation:

*"He made some mispronunciations with some of the difficult words and sometimes he stutters **but** the sound of his voice tells that he is really interested in doing his speech."*—Noelle, Sect II

*"It was evident that she was prepared. **However**, she voice could do a little tuning."*—Dino, Sect II

*"She lacked facial expression, eye contact with the audience and voice projection, **but** with the given time constraints, she was able to make use of it well."*—Kim, Sect IV

**Figure 1**  
Relative proportions of positive and negative comments



**Table 3**  
Backtracking/hedging comments for the peer

Category	Sections				TOTAL
	I	II	III	IV	
Peer	6	11	9	19	45



"She don't feel nervous when she delivered it but somehow, it seems she didn't do her best."—Elliot, Sect IV

"I think his voice was not that soft and yet not so loud. But I believe that it was loud enough for the audience to hear and understand what he was saying. His grammar was good enough though I think he said a few "ums" during the speech."—Estephanie, Sect III

The negative markers "but", "although", "however" were commonly used, while "moreover", "in addition" and "also" seemed to highlight the backtracks or hedges made by the students about their positive and/or negative comments.

The second research question is concerned with how the students rated their peers numerically and whether these numerical evaluations reflect or support their qualitative evaluations. So as not to influence the actual grades obtained by the students in the speech class, the researcher thought of using an arbitrary scale to accomplish this task. The students were asked to put a numerical equivalent of their peer evaluations on the upper right hand side of the papers they submitted. The researcher told them that this would be the composition of the scale:

It can be noted that the scale still conforms to the 70% passing mark that has been set by the department for passing students in both their writing and speech components for English One. The researcher felt that this scale did not completely deviate from the standard procedures being implemented and thus can still work as a valid numerical tool for the students to use for assessment.

Table 5 (on page 76) presents the distribution of the numerical ratings (labeled as PG [short for peer grade] ranging from 7-10, with 10 as the highest) obtained from the two special evaluators (SE, labeled as a and b) of the ten students of each class. As illustrated, the ratings fall generally within the "good" and "very good" scales, especially when peer grades match. For instance, Speakers D to F of section 3 both received a peer grade of 8, thus with their average score also being 8. This score falls under the category of "good." But the table also shows that peer evaluators' grades may not match closely with their written evaluations. For example, Speaker F in Section I received a grade of 6 from SE-a and received a grade of 9 from SE-b, giving the speaker a resulting peer average score of 7.5, which is somewhere between fair and good.

While findings for this part of the study indicate that most of the ratings given by the special evaluators did not differ significantly as most of them were close (e.g., 8 and 9, 8 and 7), speakers were relatively rated highly by their peers.

**Table 4**

Scale developed for numerical rating

Numerical Ratings and Corresponding Descriptions	
10	Excellent
9	Very Good
8	Good
7	Fair
6 and below	Poor

**Table 5**  
Numerical equivalents found in peer evaluation

Section I				Section II				Section III				Section IV			
Spkr	SE	PG	Ave.	Spkr	SE	PG	Ave.	Spkr	SE	PG	Ave.	Spkr	SE	PG	Ave.
A	a b	8 9	8.5	A	a b	6 6	6	A	a b	8 8	8	A	a b	9 9	9
B	a b	8 7	7.5	B	a b	9 8	8.5	B	a b	9 8	8.5	B	a b	9 8.7	8.85
C	a b	9 8.5	8.75	C	a b	7 8	7.5	C	a b	7 8	7.5	C	a b	10 8	9
D	a b	8 7	7.5	D	a b	6 7.5	6.75	D	a b	8 8	8	D	a b	8 8	8
E	a b	7 8	7.5	E	a b	9 10	9.5	E	a b	8 8	8	E	a b	7 8	7.5
F	a b	6 9	7.5	F	a b	7 7.5	7.25	F	a b	8 8	8	F	a b	9 9	9
G	a b	8 9	8.5	G	a b	9 9.5	9.5	G	a b	9 9	9	G	a b	8.5 8	8.25
H	a b	6 9	7.5	H	a b	8 7	7.5	H	a b	9 8	8.5	H	a b	9 8	8.5
I	a b	8 8	8	I	a b	8 7.5	7.75	I	a b	5 9	7	I	a b	9 8.5	8.75
J	a b	7 9	8	J	a b	8 9	8.5	J	a b	9 8.5	8.75	J	a b	7 8	7.5
Average			7.925	Average			7.85	Average			8.125	Average			8.435

The average grade peer grade ranged from 7.9 to 8.4. This still generally falls under "good."

As these figures appear to be consistent with the qualitative evaluations reported earlier, it can be reiterated that this high marking alludes to the students' boosting of the peer's performance rather than of the self. Further, the numerical data seem to be consistent with the patterns as found for the written evaluations that are generally positive and relatively high for the peer.

## Discussion

Results of the study may have some bearing on the characteristics of the students, which may be a function of the Filipino culture.

Peer assessment might be taken by the students to mean that a marking should be taken within the interpersonal dimension, highlighting a certain degree of politeness that supersedes the goal of the activity itself. There appears to be a

tendency to boost the ratings for the peers in the narrative report to "justify" the numerical points given. The researcher considers this finding relative to Filipino culture in terms of the Filipino value "*pakikisama*" or a person's attempt at maintaining good interpersonal relationships with others. In order to avoid negative face, positive indications to the peer are preferred. This aspect of the results seems to support Sengupta's (1998) findings that peers take the outputs of the peers positively.

A second issue that may be addressed is about the sincerity of the comments. As there was no provision in the study to verify whether the students meant what they had written (perhaps through the researcher's own initiative to interview them on their critiques or to allow them a peer-feedback session or conferencing similar to what is being done in the writing classes), it will be very difficult to substantiate that the written comments are truly reflective of their observations and assessment at the time of the performance, considering that they were written at some later point prior to the date of submission specified. In addition, the researcher also thought about the effects of this written evaluation, of whether it actually helped the students realize their own strengths and weaknesses as writers because they might have problems in expressing themselves on paper to accurately describe their evaluation. It is suggested that another study to address these concerns be undertaken in the future to the Filipino speech students of English One.

The researcher also sees some parallelisms with the observations of Cosh (1999) that rather than treating this exercise as a means of allowing the students to reflect on their own strengths and weaknesses as speakers, it was a means to criticize others. Peer assessment seems in this regard has not gone far from the usual concept of judging how others have performed. There was no indication in the students' comments that show how the performances of the peer have helped them realize their own learning in so far as public speaking is concerned. Most of the comments for the self do acknowledge their weaknesses, but they also come from the point of view of criticism, rather than through realization or introspection as an offshoot of the modeling made by others in the class. As mentioned earlier, the peers have a tendency to boost the speakers and downplay their own achievements.

This may seem speculative however, without the verification from the students themselves if such is indeed the case, but there is still a possibility that the students treated this exercise as simply a "written requirement" that they had to accomplish as far as pleasing the teacher is concerned.

This aspect of the study relates to metacognition—that the students have yet to be fully aware of their strengths and weaknesses as learners, to learn to set more specific goals as far as their own performances in the class are concerned, or even to set goals in helping their peers through the assessment process, so as to maximize what the class has to offer. On the part of the speech class teacher, more training may be needed to provide for students to utilize their metacognitive skills. When such training has taken place, it is believed to remedy the "mechanical" nature of peer assessment as simply a requirement in the class, but a collaborative endeavor for both peers and evaluators.

One thing is certain, however: that the study was able to at least for a time, allow the students to write their thoughts, as a way of testing their listening and recall skills as audience of the performances.

Finally, a third issue is whether the format of the assessment tool used (the standard rubric and the parameters given by the teacher as a guide) is the most appropriate feedback tool given the goals of the researcher in her class. There might be a more appropriate system or evaluation tool that can be used in the Filipino speech classroom that will minimize, if not eliminate, the researcher's observation that the students were merely being polite to her and the peers that are being evaluated. One example of a support activity would be through semi-structured interviews (as suggested by Tsui & Ng, 2000) that would allow evaluators and speakers to recall what they did and to properly frame their comments.

Corollary to this is the learning element involved in peer assessment. As a practice, the researcher distributed an informal evaluation form to serve as an exit interview for the students enrolled in all her classes at the end of the term. This served to give her feedback about the strengths of the class and what the students thought could be improved for future classes. Although this exercise was not part of the current research, it yielded some information about the students' perceptions of peer assessment, which are worth mentioning as they are related to the issue being explored.

One student indicated in this end-of-the-term evaluation that one of the things he learned from the class was that there was a structure involved in the assessment process and that evaluating a speech was not a case of grading based on "gut feelings." Some said that they were able to express comments on paper without having to offend the speaker by being too confrontational. Others indicated that the presence of the speech band guided them during the evaluation as it contained behaviors/descriptors that characterized the level of performance. Another student said that she learned more about herself as a speaker because of the activities conducted in the class. She believed that the evaluation component taught her to change her initial impression that being in a speech class meant that the teacher "judged" the students based on the errors that they committed, but rather, it made her more conscious about her speech as well by articulating what she thought of the others.

Another point raised in the evaluation was the point of being assigned as special evaluator. Some students felt that it gave them a chance to be "more involved" in their peers performance, rather than by being mere an audience. For them, it was a means to "help improve" their classmates' speeches.

Because the students were required to write down their opinions in contrast with expressing them orally, this writing process gave them an opportunity to reflect on what they had done and witnessed during the actual speech event. The efforts of the students should not be misconstrued as being mechanical or perfunctory as the mere exercise of providing them an opportunity of giving feedback allowed them the opportunity to express their ideas in a way that would not intimidate them so much as in situations where they were tasked to orally express their opinions in front of the whole class on a peer's performance

immediately after the said performance. In the latter case, more politeness or hedges, attenuating the real ideas of the evaluator, may be expected. Writing down their feedback also allowed them to "participate" in the speech production process, albeit as spectators or audiences. Opportunities such as this would still be considered better off than simply letting the students sit inside the class staring at their classmates blankly.

### Implications

An implication of the study relates to "positioning" and voice (cf. Ivanic & Camps, 2001). The students may be trapped at the "ideational positioning" level which is talking or writing about *something* rather than positioning themselves interpersonally, that is to say, doing the talking or writing to address the peer or the interlocutor. The ideational positioning may help explain their beliefs and values as well as its cultural underpinnings on the way they commented and on the way they viewed the exercise, the peer and even the teacher. The comments for the peer attest to this pattern of commenting on some of the observed negative behaviors done by the peer only to be taken back through the follow-up positive comments. As stated previously, more training is needed to improve the way students give feedback, and to satisfy the issues related to voice and positioning, the teacher may need to strengthen students' metacognitive skills as these skills are related to speech and peer evaluation.

A second implication of the study bears relevance to the teacher. Filipino L2 speech teachers may need to be more aware of the value of peer assessment that goes beyond its function as a requirement in class. Filipino L2 speech teachers may need to be more sensitive to the actual comments made on paper, or to its implications (e.g. on the attitudes of students toward themselves as speakers or to their peers, and their reporting styles, and their cultural underpinnings). Even the researcher is not exempted from noting the above comments. These issues could be further explored to improve the instruction of speech in the Filipino L2 context, and to strengthen assessment, not only for the teachers' benefit, but more so for the benefits of students themselves.

The third implication relates to culture. Filipino L2 learners have yet to be fully trained on assessment to the extent that both the qualitative and quantitative aspects of assessing peers' speech performance would be more reflective of cultural beliefs on what constitutes "good speech" rather than what the teacher is perceived to be expecting from them. Perhaps, the Filipino DLSU speech teacher may at the start of classes, discuss a variety of these issues to the students (assessment procedures, expectations, desirable traits during the performance) a bit beyond what is included in the usual syllabus being distributed. By doing so, students would be more appreciative of all the types of activities and participation being expected of them, making them to appreciate speech classes more, rather than simply treating them as a requirement in their respective undergraduate programs. This way, their true "voices" may be heard.

### Summary and conclusion

It was argued that written peer evaluation in a speech class is as important an area of investigation as that of the peer evaluation done in a writing class. This was demonstrated in the study conducted on the written peer evaluations of Filipino DLSU English One students' descriptive and demonstrative speeches. Instead of using the writing band as a basis for evaluating quality of their comments, the speech band was utilized in looking at the areas of speech where the students had to frame their evaluation within a separate narrative critique submitted to the teacher.

The paper attempted to describe how the students viewed and evaluated their peers' performance by looking at how they wrote their comments and determined whether these peer narrative evaluations supported the numerical evaluations given for the peer. Findings of the study indicated that students mostly evaluated through a comparison and contrast of their positive and negative impressions of the peers' performances. Positive comments for the peer generally focused on the following items: content, organization and vocal expression, while negative comments for the peer mainly focused on the following elements: the categories of physical expression, vocal expression and others. Findings of the study also showed that students offered a more positive opinion of their peers' performances. Finally, the numerical ratings given are generally positive as supported by the summary statistics presented.

The study recommends that this be replicated by incorporating some improvements in the methodology. One improvement may be the inclusion of a reflective task so as to tie this up with any pre- and post-testing that will be made. Such a reflective task would give both the teacher and students validity in the comments made on the performance, which was a limitation in the current study. A suggestion would be to include oral commenting as a follow-up to allow student authors to negotiate the meaning with their peers based on the written comments.

Future studies could also look at incorporating more metacognitive training so that students would be more sensitive to the learning processes and goals of the class. The positive and negative categories used in the study may be further refined to capture other points of interest that may be raised.

In addition, although the study was somehow able to show that students can verbalize their opinions on paper, they still could not directly relate those comments to their numerical equivalents. Students need to be further trained on providing evaluation or feedback that is more accurate or more representative of what the teacher also sees in the performance. With this in mind, a study on improving assessment skills could be done. Furthermore, the speech band that the Filipino DLSU English One Students used is a standard band that has been developed using learners from other cultures (mostly western). The researcher believes that it may be high time for the university faculty members who are speech teachers be given a chance to develop a band more suitable to the culture of its L2 learners, without deviating too much on the standard parameters of speech assessment. Other learning institutions in the country that also use foreign rubrics may opt to devise more "local" bands to fit what both students and the teacher perceive as their "classroom standard."

To conclude, the study was able to provide an insight into how the narrative peer evaluation may be used in a L2 speech classroom. This type of evaluation offered an alternative to using just the numerical evaluation given by the teachers at DLSU for their lack of sufficient time to openly and extensively discuss performances in class. The written peer evaluation allowed the students to "speak" on matters that were otherwise "unspeakable" in class, even if it initially meant addressing the teacher, rather than the peers themselves.

### Acknowledgements

The researcher would like to thank the DLSU College of Education for providing funding for this research and for Ms. Elaine Llarena who served as research assistant.

### References

- Ashwell, T. (2000). Patterns of teacher response to student writing in a multiple-draft composition classroom: Is content feedback followed by form feedback the best method? *Journal of Second Language Writing*, 9(3), 227-257.
- Boud, D.J., & Tyree, A.L. (1979). Self and peer assessment in professional education: A preliminary study in law. *Journal of Society of Public Teachers of Law*, 15, 65-74.
- Cosh, J. (1999). Peer observation: A reflective. *ELT Journal*, 53(1), 22-27.
- Heilenman, K.L. (1990). Self-assessment of second language ability: the role of response effects. *Language Testing*, 7, 174-201.
- Ivanic, R., & Camps, D. (2001). I am how I sound: Voice as self-representation in L2 writing. *Journal of Second Language Writing*, 10(1-2), 3-33.
- Mangelsdorf, K. (1992). Peer reviews in the ESL composition classroom: What do the students think? *ELT Journal*, 46(3), 274-283.
- Northern Ireland Curriculum. (n.d.). Retrieved on February 17, 2004 <http://www.ccea.org.uk/ks3/pdf/book2/bk2sect3.pdf>
- Patri, M. (2002). The influence of peer feedback on self and peer assessment of oral skills. *Language Testing*, 19(2), 109-131.
- Ross, S. (1998). Self-assessment in second language testing: A meta-analysis and analysis of experiential factors. *Language Testing*, 15(1), 1-20.
- Sengupta, S. (1998). Peer evaluation: "I am not the teacher." *ELT Journal*, 52(1), 19-27.
- Shore, T. and Shore, L. (1992). Construct validity of self- and peer evaluations of performance dimensions in an Assessment Center. *Journal of Applied Psychology*, 27(1), 42-54.
- Stefani, L.A.J. (1994). Peer, self and tutor assessment: Relative reliabilities. *Studies in Higher Education*, 19, 69-75.
- Tsui, A., & Ng M. (2000). Do secondary L2 writers benefit from peer comments? *Journal of Second Language Writing*, 9(2), 147-170.
- Williams, E. (1992). Student attitudes towards approaches to learning and assessment. *Assessment and Evaluation in Higher Education*, 17, 45-58.



### Appendix: Speech Band

\* adapted from Clayton College and State University Communications 1001 and 1002 Version, August 1998/band may be modified for group or impromptu speaking

CRITERIA	EXCELLENT	GOOD	FAIR	UNSATISFACTORY
Information Base Appropriate for the Presentation	The speaker demonstrates a thoroughly comprehensive and sophisticated information base.	The speaker demonstrates a somewhat comprehensive and sophisticated information base.	The speaker demonstrates an information base barely sufficient for accomplishing the assigned task.	The speaker is unable to demonstrate an acceptable information base for accomplishing the assigned task.
Content Appropriate to Purpose, Listener and Occasion	<b>Development of content:</b> The speaker effectively limits and focuses the topics, selects, adapts and develops the content. The content is interesting, clear and appropriate to frame of reference of the listener. <b>Use of visual aids:</b> Appropriate visual aids contribute to the listener's understanding of the content.	<b>Development of content:</b> The speaker sufficiently limits and focuses the topic, selects, adapts and develops content. The content is generally relevant, reliable and sufficient. <b>Use of visual aids:</b> Appropriate visual aids contribute to the listener's understanding of the content.	<b>Development of content:</b> The speaker has not appropriately limited and focused the topic, resulting in inadequate development as it relates to speaker's purpose. The speaker attempts to provide supporting material, but material is lacking in either quality, variety, accuracy or relevance. <b>Use of visual aids:</b> Visual aids are either inappropriate or ineffectively used.	<b>Development of content:</b> The speaker has inappropriately limited the topic. The content does not help the speaker achieve the specific purpose. The supporting details are for the most part either unclear, irrelevant, unreliable, inadequate, insufficient or clearly inaccurate. <b>Use of visual aids:</b> Visual aids are either inappropriate, insufficiently used or lacking at all.
Organization Appropriate to Purpose, Listener and Occasion	In the intro, the speaker comfortably gains favorable attention and motivation for listening. The central idea was effectively focused on the material. The body is a clear unit if distinct major supporting ideas with specific details. Conclusion is effectively suited to speaker's purpose and to other elements of presentation. Transitions are smooth providing the listener with a sense of continuity for the entire presentation.	The opening comments reveal a strategy designed to gain favorable attention and establishes motivation for listening. The speaker has a central idea, the direction of ideas is established, the body has distinct major ideas with specific details, the concluding comments are appropriately related to purpose. The order of introductory or concluding comments and order of major supporting ideas were effective in achieving purpose. The speaker provides clear but somewhat stilted transition between elements of presentation.	The speaker has difficulty introducing the subject, first few comments somewhat unrelated, weak or unimaginative. Central idea is flawed, either the attempt to focus the info reveals topic but not central idea or the central idea should have been better placed. The listener finds difficulty recognizing the body of presentation and the supporting ideas. There is difficulty ending the presentation, concluding comments inappropriate or inadequate for purpose. Transitions are awkward or weak providing little continuity for the listener.	The opening remarks are either inappropriate to the presentation, or unlikely to motivate listening. There is no focusing central idea, or placing is inappropriate for purpose of speech. Body of presentation lacks major supporting ideas. The speaker fails to provide a closing statement. The major and supporting ideas confuse the listener. The speaker has no clear transitions between elements of presentation to provide continuity.
Physical Expression Appropriate for	The speaker looks genuinely interested in communicating, the facial expression enhances the verbal message and give	The speaker shows interest, facial expressions are consistent with speaker's voice, language and content	The speaker has limited facial expressions. Distracting behaviors or the absence of appropriate behaviors is	The speaker lacks facial expressions. Distracting behaviors or the absence of appropriate behaviors obstructs

the Listener and Occasion

appropriate emphasis to the content. Body movements, posture and gestures are

of message. Distracting behaviors or absence of appropriate behaviors do

noticeable and occasionally interferes with communication. Appearance does

communication. The appearance distracts from the message. Lack of eye



<p>Expression Appropriate for</p> <p>communicating, the facial expression enhances the verbal message and give</p> <p>expressions are consistent with speaker's voice, language and content</p> <p>expressions. Distracting behaviors or the absence of appropriate behaviors is</p> <p>Distracting behaviors or the absence of appropriate behaviors obstructs</p>	<p>the Listener and Occasion</p> <p>appropriate emphasis to the content. Body movements, posture and gestures are natural and relaxed to reinforce the message and give emphasis to content. The speaker's appearance is tailored to the occasion and expectations of the listener. Eye contact appears to be varied, purposeful and comfortable.</p>	<p>of message. Distracting behaviors or absence of appropriate behaviors do not interfere with communication. Appearance is appropriate for occasion and listener. Eye contact is consistently maintained.</p>	<p>noticeable and occasionally interferes with communication. Appearance does not distract from the message. Eye contact is inconsistent.</p>	<p>communication. The appearance distracts from the message. Lack of eye contact is obstructing to the information exchange.</p>
<p>Vocal Delivery Appropriate for the Listener and Occasion</p>	<p>Volume is appropriate for the listener and occasion. Articulation is consistently distinct and clear. Presentation free of mispronunciations. Speaker sounds genuinely interested in the topic by using effective vocal delivery—varying pitch, rate and volume throughout the presentation. Pauses are used to enhance impact of message, notes if used are used effectively. An occasional vocalized pause may occur but not distracting.</p>	<p>Volume is generally appropriate to occasion and listener, articulation distinct and clear, pronunciation for most part is correct but an occasional lapse is noticeable. The speaker sounds interested in the topic and shows some vocal variety. Notes may assist but do not interfere with delivery. Vocalized pauses cause occasional minor distractions.</p>	<p>Volume is generally too loud or too soft for the listener and occasion. Articulation is hampered by occasional lazy, slurred or chopped enunciation. Pronunciation interferes with listener's understanding of the message. Speaker lacks vocal variety thus having somewhat monotonous delivery which hamper's the listener's reception of the message. Moments of apparent reciting, reading, or referring to notes occasionally interferes with delivery. Vocalized pauses are distracting.</p>	<p>Volume is generally too loud or too soft for the listener and occasion. Poor articulation and mispronunciation obstructs communication with listener. Reciting, reading or constant referral to notes adversely affects vocal delivery. The speaker sounds uninterested or bored with the message in communicating successfully with listener. Vocalized pauses obstruct communication with the listener.</p>
<p>Language Appropriate for the Listener and Occasion</p>	<p>The speaker makes no noticeable grammatical errors which interfere with perceived credibility or with the listener's understanding of the message. Any deviation from standard spoken English seems to be chosen for an appropriate reason. The speaker's dialect should not interfere with the communication process. The speaker chooses language that is clear, direct, concrete, vivid and appropriate. Figurative language and imagery indicate a measure of originality and sophistication in speakers control of the language. Jargon if used is appropriate for listener. Sentence structure is varied for emphasis and stylistic impact appropriate to both listener and occasion.</p>	<p>The speaker shows general mastery of language standards throughout communication. Deviation from standard spoken English seems to be chosen for an appropriate reason. Language is generally clear, direct, concrete and appropriate. Expressions are free of slang, colloquialisms or idioms that are inappropriate for listener. If used, jargon is appropriate to listener.</p>	<p>Grammatical errors are noticeable, sometimes distracting, and at times, inappropriate for the language background of listener. Style sometimes interferes with the listening process: either speaker uses language that is vague or ambiguous or empty, shows awkward syntax, or language that contains expressions and syntax that is more appropriate for written style.</p>	<p>Frequent errors in standard spoken English are noticeably objectionable and make it difficult for the listener to focus on the message. Level of standard English used is consistently inappropriate for the listener. The style interferes with the listening process: the speaker either uses language that is vague, ambiguous or empty, or has awkward syntax or the speaker relies wholly on written style for the presentation of the message.</p>

