Teaching and Assessing Speaking Performance through Analytic Scoring Approach

Luu Trong Tuan

University for Natural Resources and Environment, Ho Chi Minh City, Vietnam Email: luutrongtuan@vnn.vn

Abstract—The research sought to discern if the analytic scoring approach would be more effective for improving EFL learners' speaking performance. The two groups, analytic group and holistic group, were compared in terms of the test results from analytic scoring approach and holistic scoring approach respectively. The study also found the students' positive attitude towards the adoption of analytic scoring approach in teaching and assessing speaking skill.

Index Terms-speaking, analytic scoring approach, holistic scoring approach, EFL

I. INTRODUCTION

The holistic scoring approach has been the foundation for teaching and assessing speaking performance of the EFL students at University for Natural Resources and Environment, Ho Chi Minh City. Nonetheless, the holistic scoring approach in teaching and assessment does not provide useful anatomical guidance for students to improve their speaking competence. Obtaining the score for speaking, students do not know why they succeed or fail since a single score does not allow [students] to distinguish between various aspects of speaking such as control of syntax, depth of vocabulary, organization, and so on.

Bachman and Palmer (1996) suggested a framework for testing in terms of the usefulness. The framework can be relevant in helping teachers decide which type of test to use. This framework proposed six qualities of test usefulness: Reliability, Construct Validity, Authenticity, Inter-activeness, Impact, and Practicality (Bachman and Palmer, 1996, pp. 17-38). Weigle (2002) commented on Bachman and Palmer's framework, showing a comparison of holistic and analytic scales based on the same six qualities of test usefulness. On reliability, the analytic scales will be more reliable than the holistic scales even though the holistic scoring scales are acceptable. On construct validity, the analytic scales are more appropriate to second language students. On impact, the analytic scales will provide students with more information about their competence; teachers and educators also are benefited with the analytic scales will be more disadvantageous than the holistic scales when practicality is taken into consideration since it is a time-consuming and expensive method of scoring.

The aim of this research was to examine if the analytic scoring approach in teaching and assessment would be more effective for improving students' speaking performance. The research also explored the students' attitude towards the implementation of analytic scoring approach in speaking teaching and assessment. The research question guiding this research thus encompasses:

1. Does teaching and assessing speaking performance analytically improve EFL students' speaking competence better than teaching and assessing speaking performance holistically?

2. How do EFL students reflect on the implementation of the analytic scoring approach in speaking teaching and assessment?

II. LITERATURE REVIEW

Analytic Scoring Schemes

Analytic scoring schemes are means of assessment by breaking down the objective of final products into criteria parts, and each part is scored independently. The procedures of this method involve the separation of the various features of a discourse into categories for scoring purpose (Park, 2004, p. 1). The total score is the sum of the rating for all of the parts that are being evaluated. When using analytic scoring schemes, it is necessary to treat each criterion or part as separate to avoid bias towards the whole product. Depending on the purpose of the assessment, speaking performance might be rated on such criteria as content, organization, cohesion, register, vocabulary, grammar, or mechanics. Contrary to the holistic scoring scheme, this method of scoring avoids the potential flaw in global impression band scales of uneven development in the different criteria (Weir, 2005, p. 189). Moreover, with this approach of assessment, a teacher is easy to give a higher scoring for a certain criterion by giving a certain coefficient when he/she think that his/her students should focus on the criterion. For example, if a teacher pays more attention to the organization of the

speech, he/she can give a coefficient of two to the criteria before the total score of the speaking performance is calculated.

Advantages of Analytic Scoring Schemes

Analytic scoring schemes are preferred over holistic schemes by many speaking specialists for a number of reasons. First, as mentioned above, it provides more useful diagnostic information about students' speaking abilities. That is, it tells learners where their weaknesses are and where their strengths are. Analytic scoring has been considered as more interpretable scoring approach because it accesses the examinee's specific strengths and weaknesses and identifies the particular components of speaking discourse that an examinee needs to develop (Downing and Haladyna, 2006, p. 314). Holistic scores provide valuable information for an overall categorization of speaking ability, but analytic scores provide more diagnostic information. The information also allows instructors and curriculum developers to tailor instructions more closely to the needs of their students. To a certain extent, the analytic scoring scale is a useful tool to provide teachers' feedback for students on the areas of students' strengths and weaknesses. Park (2004, p. 2) pinpointed that the explicitness of the analytic scoring scheme guides offers teachers a potentially valuable tool for providing speakers with consistent and direct feedback. Second, analytic scoring schemes are particularly useful for second language learners, who are more likely to show a marked or uneven profile across different aspects of speaking discourse. Some second language learners may have excellent speaking skill in terms of content and organization, but may have much lower grammatical control; others may have an excellent control of sentence structure, but may not know how to organize their speech in a logical way. On this aspect, the analytic scoring scales can show students that they have made progress over time in some or all dimensions when the same rubric categories are used repeatedly (Moskal, 2000).

Analytic schemes have also been found to be particularly useful for scorers who are relatively inexperienced (Weir, 2005, p. 190). Weir in his 1990 research reported that a multi-trait analytic mark scheme is seen as a useful tool for the training and standardization of new examiners (Weir, 2005, p. 190). Other authors maintained that, compared to holistic scoring schemes, analytic scoring schemes are easier to train scorers to use it, an inexperienced scorer may find it easier to work with an analytic scoring scheme than a holistic scoring one because they can evaluate specific textual criteria (Park, 2004, p. 2; McNamara, 1996). Thus, inexperienced scorers may find it easier to work with an analytic scale than a holistic one.

Disadvantages of Analytic Scoring Schemes

The major disadvantage of scoring analytically is that it takes a lot of time to rate speaking performance since examiners are required to make more than one decision for every speaking performance. When scoring analytically, an examiner has to check, consider, and score each criterion of the speaking ability and then gives a total score depending on the coefficient put forward.

Critics of analytic scoring schemes also point out that measuring the quality of a text by tallying accumulated subskill scores diminishes the interconnectedness of spoken discourse. At this aspect, it is thought that "the whole should be greater than the sum of its part". Measuring the quality of a spoken discourse by tallying accumulated sub-skill gives the false impression that speaking can be understood and fairly assessed by analyzing autonomous discourse features (Park, 2004, p. 3). Hughes (1989) pinpointed that concentration on the different aspects may divert attention from the overall effect of the speech. Inasmuch as the whole is often greater than the sum of its parts, a composite score may be very reliable but not valid (Hughes 1989, pp. 93-94). In this aspect, the analytic scoring often has the tendencies to reduce and oversimplify the components of speaking, and to emphasize the flaws than the strengths of speaking.

Hughes (2003) warned that in scoring analytically, the criterion scored first may affect on subsequent criteria which are scored later, making the overall effect of a speech diverted to an individual criterion. Futcher (2009), based on Thorndike's idea, named this phenomenon as the Halo Effect of the analytic scoring. Thorndike in 1920 defined the phenomenon as *a problem that arises in data collection when there is carry-over from one judgment to another*. In other words, when scorers are asked to make multiple judgments they really make one, and this affects all other judgments. If scorers are give five scales each with nine points, and they award a score of five on the first scale for a speech, it is highly likely that they will score five on the second and subsequent scale, and be extremely reluctant to move too far away from this generally. As a result what we find is that profiles tend to be "flat", defeating the aim of providing informative, rich information, on learner performance (Futcher, 2009). Consequently, criteria scales may not be used effectively according to their internal criteria, resulting in a halo effect in which one criteria score may influence another.

An additional problem with some analytic scoring schemes is that even experienced essay judges sometimes find it difficult to assign numerical scores based on certain descriptors (Hamp-Lyons, 1989). In this aspect, there are possibilities for scorers to disagree with one another. It is more difficult to achieve intra- and inter-rater reliability on all of the dimensions in an analytic scoring scheme than on a single score yielded by a holistic scale. Also, on the scorers' part, McNamara (1996) exposed that there are some evidences proving that scorers tend to evaluate grammar-related categories more harshly than they do other categories (McNamara, 1996), thereby overemphasizing the role of accuracy in providing a profile of students' proficiency. This disadvantage is inevitable, especially with un-trained or un-experienced scorers. Grammar-related categories are somewhat wrong – right categories whereas other categories are judgments. Focusing on wrong – right categories will always be easier than judgments. White (1985) added other limits

of analytic scoring. That is the lack of agreement about what separate traits exist and its tendency to complicate the assignment of the scores for listeners, increasing time and therefore costs.

III. METHODOLOGY

A. Participants

Students

The two classes comprising 104 students at University for Natural Resources and Environment, Ho Chi Minh City were invited to participate in the research as the experimental group (the analytic group) in which the analytic scoring approach was utilized for teaching and assessing speaking and the control group (the holistic group) in which the holistic scoring approach was utilized for teaching and assessing speaking.

The analytic group consisted of 51 students, among which 38 students were female, accounting for 74.51% and 13 students were male, accounting for 25.49%. This class was chosen to be the analytic group in this research since the group's mean of the pretest was 6.78, lower than 6.81 of the holistic group. This choice would give the study more reliable when the holistic group seemed to be better than the analytic group at the beginning of the research.

The holistic group consisted of 53 students, among which 40 students were female, accounting for 75.47% and 13 students were male, accounting for 24.53%.

Teacher

A female teacher was invited to teach the two groups to avoid the researcher's bias. The teacher holds a master degree in TESOL and has over six years of experience in teaching speaking.

B. Instruments

Pretest and posttest

The pretest was the final test that the students in the two groups took for the subject of speaking in their previous semester. The score of the posttest was the aggregation of the scores from six speaking tests through the speaking course.

The analytic scoring scale for this study was adapted from the scales suggested by Nakamura (2004, p. 47) and Hughey *et al.* (1983, p. 140). The Nakamura scoring scale is on the scale of four for each criterion among five criteria of Originality of Content, Organization, Vocabulary, Grammar and Logical Consistency. The Hughey *et al.*'s scale has five criteria of Content, Organization, Vocabulary, Language use and Mechanic. The analytic scoring scale used as the basis for teaching and assessing speaking in this research contains five criteria: 1) Coherence, 2) Content, 3) Grammar and Structure, 4) Language used (consisting of Vocabulary, Spelling and Word used), and 5) Organization.

Questionnaire

The questionnaire used in the study was a six-scale Likert-type one in Vietnamese language. The six-scale response was used in the questionnaire to prevent respondents from choosing a "sit the fence" attitude by making the most neutral possible answer (Brown, 2000, p. 1). The type was designed to ask respondents to choose either in the positive or negative attitude towards the use of the analytic scoring approach.

The questionnaire was used for investigating students' attitudes towards the use of the analytic scoring approach in teaching and assessing their speaking performance. The questionnaire has six items.

- Item 1 (*The analytic scoring approach helps students enhance their speaking skill*) was designed to examine students' attitudes towards the role of the analytic scoring approach in enhancing students' speaking competence. This item has the response categories of Agreement. The six options for the response categories are *Strongly disagreed – Disagreed – Somewhat disagreed – Somewhat agreed – Agreed – Strongly agreed*.

- Item 2 (*The analytic scoring approach is essential for students to enhance their speaking competence*) was designed to examine students' attitudes towards the necessity of using the analytic scoring approach. This item has the response categories of Agreement. The six options for the response categories are *Strongly disagreed – Disagreed – Somewhat disagreed – Somewhat agreed – Agreed – Strongly agreed*.

- Item 3 (*I am interested in discerning my strengths and weaknesses in speaking skill*) was designed to examine students' attitudes towards the fact that the analytic scoring approach displays students' strengths and weaknesses in speaking skill. This item has the response categories of Agreement. This item was expected to be correlated with Item 1. The six options for the response categories are *Strongly disagreed – Disagreed – Somewhat disagreed – Somewhat agreed – Agreed – Strongly agreed*.

- Item 4 (*How often did you apply the analytic scoring approach to their learning strategies outside the classroom?*) was designed to examine how students applied the analytic scoring approach to their learning strategies. This item has the response categories of Frequency. The options for the responding categories are *Very frequently – Frequently – Occasionally – Rarely – Very rarely – Never*.

- Item 5 (*Do you want the analytic scoring approach to be applied to your other speaking courses*) was designed to examine students' expectations for the continuous application of the analytic scoring approach to their learning of speaking. This item has the response categories of Likelihood. The options for the responding categories are *To a great extent – Probably – Somewhat – Possibly – Very little – Not at all.*

- Item 6 (*I am pleased with my improvement in the speaking competence after attending the course*) was designed to examine students' content with their improvement on the speaking competence after attending the course. This item has the response categories of Likelihood. The options for the responding categories are *Definitely – Very probably – Probably – Probably – Probably not*.

IV. FINDINGS AND DISCUSSIONS

A. Students in the Analytic Group Enhanced Their Speaking Performance Better than Students in the Holistic Group during the Process of Analytic Teaching and Assessment

Table 1 and Table 2 recapitulate the results that the two groups obtained through the six speaking tests in the study. Table 1 is for the holistic group and Table 2 is for the analytic group. For the columns of Mean, Min and Max in the tables, the students in the holistic group did not improve much during the course from the first speaking test to the last one. These students earned averagely 6.31 for speaking test 1 but they averagely earned only 6.58 for speaking test 6. They improved with their minimum and maximum scores by one mark for each parameter; concretely from 3.50 up to 4.50 for their minimum scores and from 8.00 up to 9.00 for their maximum. It is also noted that the students in the holistic group obtained relatively static results from speaking test 1 to speaking test 6 when the means are analyzed. The highest value of the means is for speaking test 6 (6.58) and the lowest one is for speaking test 1 (6.31).

			IADLE I.				
	SUMMARY OF TH	E RESULTS OF T	HE HOLISTIC GF	ROUP THROUGH SIX SPE	AKING TESTS		
Holistic group							
				Difference in Mec	lian -		
	Mean	Median	Range	Mean	Min	Max	
Speaking test 1	6.31	6.00	4.00	(-) .32*	3.50	8.00	
Speaking test 2	6.38	6.50	4.50	.13	3.50	8.00	
Speaking test 3	6.37	6.50	4.50	.15	4.00	8.50	
Speaking test 4	6.48	6.50	4.00	.03	4.50	8.50	
Speaking test 5	6.50	6.50	4.50	.10	4.50	8.00	
Speaking test 6	6 58	6 50	4 50	03	4 50	9.00	

TABLE 1

(-)* indicates that the mean is higher than the median

For the students in the analytic group, the statistic summary denotes an improvement during the process of learning. For the means, these students improved from 6.33 for speaking test 1 to 7.06 for speaking test 6. These improvements are for the four last speaking tests with 6.68, 7.03, 6.93 and 7.06, while the means for the two first speaking tests are stable with 6.33 and 6.35. The differences between the means and the medians also denote a prominence for the analytic group, especially with the four last speaking tests. Speaking tests 3 and 4 which have negative results prove that the number of students obtained scores from the median and above is higher than the number of students obtained scores in the two sets are spread all over the set from the minimum score to the maximum one.

SUMMARY OF THE RESULTS OF THE ANALYTIC GROUP THROUGH SIX SPEAKING TESTS							
Analytic group							
	Difference in Median -						
	Mean	Median	Range	Mean	Min	Max	
Speaking test 1	6.33	6.50	4.00	.17	4.50	8.50	
Speaking test 2	6.35	6.50	5.00	.15	4.00	9.00	
Speaking test 3	6.68	6.50	3.50	(-) .18*	5.50	9.00	
Speaking test 4	7.03	6.50	4.00	(-) .53*	5.50	9.50	
Speaking test 5	6.93	7.00	4.50	.07	5.00	9.50	
Speaking test 6	7.06	7.00	4.50	.01	4.50	9.00	

TABLE 2.

(-)* indicates that the mean is higher than the median

The comparison between Table 1 for the holistic group and Table 2 for the analytic group substatiates that the analytic group obtained better results than the holistic group when the process of teaching speaking analytically. The improvement of the analytic group can be attributed to the use of analytic approach in informing students' strengths and weaknesses. After the first two speaking tests, the students in the analytic groups perceived their strengths and weaknesses in their speaking skill. This awareness helped them adjust their learning strategies to surmount their weaknesses. As a result, these students improved their speaking performance better than students in the holistic groups on the last four speaking tests.

The t-test results also display the differences between the two groups' score sets through the process. For the frist three speaking tests, the differences are *not statistically significant*. On the contrary, all of the differences between the two groups' score sets of the last three speaking tests are *very statistically significant* for speaking tests 4 and 5 and *extremely statistically significant* for speaking test 6. The p values for these t-tests are extremely small, .00112, .00301, .00059, respectively. It should be noted that these values are smaller and smaller from speaking

test 4 to speaking test 6, denoting that the differences are more and more statiscally significant. Table 3 summarizes the results of the t-tests for the six speaking tests.

T-TEST RESULTS FOR THE SIX SPEAKING TESTS							
	Means		T-test results				
	Analyti c group	Holisti c group	two-tailed t _{statistical}	df	P value	Conclusion	
Speaking test 1	6.33	6.31	.80882	102	.420	not statistically significant	
Speaking test 2	6.35	6.38	.18638	102	.852	not statistically significant	
Speaking test 3	6.68	6.37	1.07373	102	.285	not statistically significant	
Speaking test 4	7.03	6.48	3.35399	102	.00112	very statistically significant	
Speaking test 5	6.93	6.50	3.03913	102	.00301	very statistically significant	
Speaking test 6	7.06	6.58	3.54494	102	.00059	extremely statistically significant	

TABLE 3

This finding is in accordance with Weigle's comment on Bachman and Palmer's test usefulness of Construct Validity. Weigle (2002) claimed that analytic scoring approach is more appropriate for second language speakers as different aspects of speaking performance develop at different rate. This idea is also shared by other researchers (Downing and Haladyna, 2006; Park, 2004; Moskal, 2000). Especially, Moskal (2002) and Park (2004) claimed that the analytic scoring schemes would be useful for the process of learning speaking skill. During the process of the course, it was noted that the students in the analytic groups improved their speaking competence better than those in the holistic group through each speaking test, especially the last four tests.

B. The Improvements of Students in the Analytic Group in Terms of the Five Criteria of the Analytic Assessment

The improvement of the students in the analytic group on the speaking performance can be further analyzed through their improvements on the five criteria of the analytic assessment. Table 4 displays the average scores of the analytic group based on the five analytic criteria. These scores are converted from the scale of four to the scale of ten by multiplying these scores with the coefficient of 2.5. This conversion is made to have the score in line with the current scale used at University for Natural Resources and Environment, Ho Chi Minh City.

On the whole, as Table 4 displays, the students in the analytic group made their improvements on all of five criteria. The highest improvement among the five criteria is Content. This criterion has 5.80 for the lowest score and 7.11 for the highest score, equal to 13.07%. The smallest difference is for Organization which has 6.50 for the lowest and 7.01 for the highest score. The difference is .51, equal to 5.07%.

The lowest scores for each criterion are with the first two speaking tests. Speaking test 1 has the lowest scores for Coherence, Grammar and Organization. Speaking test 2 has the lowest scores for Content and Language used. The highest scores for each criterion are with the last three speaking tests. Speaking test 4 has the highest scores for Content and Language used. Speaking test 4 has the highest scores for Coherence, Grammar and Organization. The criterion of Grammar exhibited a gradual improvement with 11.09%.

TABLE 4.							
AVERAGE SCORES OF THE ANALYTIC GROUP IN TERMS OF THE FIVE CRITERIA IN THE SIX SPEAKING TEST							
	Coherence	Content	Grammar	Language used	Organization		
Speaking test 1	5.83	6.14	6.39	6.72	6.50		
Speaking test 2	6.05	5.80	6.88	6.32	6.80		
Speaking test 3	6.13	6.26	7.37	6.80	6.81		
Speaking test 4	6.36	7.11	7.29	7.45	6.91		
Speaking test 5	6.68	6.55	7.47	7.06	6.90		
Speaking test 6	6.75	6.70	7.58	7.24	7.01		

(*) Scores are converted into the scale of ten by multiplying the average scores with 2.5

C. Students in the Analytic Group Improved Their Speaking Performance Much Better than Students in the Holistic Group from the Pretest to the Posttest

The results of the pretest denote that the two groups were almost at the same speaking competence level. The holistic group showed the slightly higher mean for the pretest (6.81 versus 6.78) and the smaller range (3.50 versus 4.00) compared to the analytic group. The t-test also provides the conclusion that the difference of the means relative to the spread of the two groups' scores is not statistically significant with the p value of .877, far much higher than the critical α value. Thus, the two groups were at the same level of English speaking when the study commenced.

Table 5 compares some parameters of the two groups in the pretests and the posttests. The comparison displays that the students in the analytic group excel their counterparts in the holistic group after four-month experiment. These students earned .66 higher for the means, .75 higher for the maximum score, 1.11 higher for the minimum score. Their

range is advantageously .36 smaller. This smaller range proves that the analytic groups' score set is more focused on the mean than the holistic group's.

 TABLE 5.

 THE COMPARISON OF THE DESCRIPTIVE STATISTICS OF THE ANALYTIC GROUP AND THE HOLISTIC GROUP FOR THE PRETEST AND POSTTEST

 Pretest
 Pretest
 Posttest difference

	Pretest		Pretest	Posttest	Posttest		
	Analytic	Holistic	difference	Analytic	Holistic	Positiest difference	
Mean	6.78	6.81	03	6.93	6.27	.66	
Max	9.00	9.00	equal	9.01	8.26	.75	
Min	5.00	5.00	equal	5.25	4.14	1.11	
Range	4.00	4.00	equal	3.76	4.12	36	

* the minus sign implies the value of holistic group is higher

D. Findings from the Questionnaire Survey

The questionnaire has six items. The first five items are used for examining students' attitudes towards the adoption of the analytic scoring approach in teaching and assessing their speaking performance. The last item, Item 6, is used for asking students to express their content with the improvement on their speaking skill after attending the course.

Findings from responses to item 1 of the questionnaire

Item 1 (*The analytic scoring approach helps students enhance their speaking skill*) was designed to examine students' attitudes towards the role of the analytic scoring approach in enhancing students' speaking competence.

The pattern of the responses displays that most of the students agreed that the analytic scoring approach helps them enhance their speaking skill to different degrees of agreement. 50 students (98.04%) of the group agreed to the statement from *somewhat agreed* to *strongly agreed*. Concretely, 76.47% *strongly agreed* with the statement, 15.69% *agreed* and 5.88% *somewhat agreed*. On the negative side of the response categories, only one student, accounting for 1.96%, *somewhat disagreed* with the statement.

The pattern of the responses to Item 1 display that there is a broad consensus among the students in the analytic group on the issue that the analytic scoring approach helps them enhance their speaking competence. This attitude is in accordance with numerous researchers' standpoints (Moskal, 2000; Downing and Haladyna, 2006; Park, 2008).

Findings from responses to item 2 of the questionnaire

Item 2 (*The analytic scoring approach is essential for students to enhance their speaking competence*) was designed to examine students' attitudes towards the necessity of using the analytic scoring approach.

The pattern of the responses displays that there are non-consensus among the students in the analytic group on the issue. Twenty-one students, accounting for 41.18%, disagreed with the statement, stating that the analytic scoring approach is not indispensable for students to enhance their speaking competence. Among these students, 12 students *strongly disagreed*, ten students *disagreed* and 13 students *somewhat disagreed* to the statement. Thirty students, accounting for 58.82%, thought that the analytic scoring approach is necessary for students to improve their speaking capability. Among these students, ten students *somewhat agreed* with the statement, 13 students *agreed* and seven students *strongly agreed*.

Even though these students have the broad consensus on the previous item which states that the analytic scoring approach will help them improve their speaking competence, they almost disagree with their classmates about the necessity of the analytic scoring approach. Nevertheless, this polarization is not opposite to other researchers' views. None of the researchers have mentioned that the analytic scoring schemes are essential for students to improve their speaking capability.

Findings from responses to item 3 of the questionnaire

Item 3 (*I am interested in discerning my strengths and weaknesses in speaking skill*) was designed to examine if the information about their strengths and weaknesses in speaking skill interest them.

In response to Item 3, all of the students agreed to different degrees to the statement which states that they are interested in discerning their strengths and weakness in speaking skill. Eight students, accounting for 15.69%, responded that they *somewhat agreed* to the statement. Fifteen students, accounting for 29.41%, *agreed* to the statement. Over half of the class, 28 out of 51 students, responded that they *strongly agreed* to the statement.

This finding is, to some extent, correlated with Item 1 as it is posited. The reliability statistics display that the Cronbach alpha's for the two items is .866 and the inter-item correlation is .767. These high values are in accordance with the percentage statistics when 50 students (98.04%) of the group are on the side of agreement for Item 1 and all of the students are on the side of agreement for Item 3.

Findings from responses to item 4 of the questionnaire

Item 4 (*How often did you apply the analytic scoring approach to their learning strategies outside the classroom?*) was designed to examine how students applied the analytic scoring approach to their learning strategies.

The pattern of the responses to Item 4 displays that all of the students in the analytic group did resort to the analytic scoring approach to adapt their learning strategies outside the classroom. 37 students (72.55%) of the group responded that they often based on the analytic scoring approach to prepare their speech. Among these, 29.41% *occasionally*, 35.29% *frequently* and 7.84% *very frequently* paid attention to the scores. However, 14 students (24.45%) of the group *very rarely* to *rarely* used the analytic scoring approach in their speaking strategies.

Findings from responses to item 5 of the questionnaire

Item 5 (*Do you want the analytic scoring approach to be applied to your other speaking courses*) was designed to examine students' expectations for the continuous application of the analytic scoring approach to their learning of speaking.

Responding to Item 5, the students in the analytic group displayed a broad consensus on the issue. All of the ideas are on the positive side of the attitudes. 20 students (39.22%) are eager to have the analytic scoring scheme used for their other speaking courses. 27 students (52.94%) express that they *probably* and four students (7.84%) contend that they *somewhat* want to have the analytic scoring scheme used for their other speaking courses. None of the students gave the attitudes from *not at all* to *possibly*. The students' prominent favor proves that the students understood that the analytic scoring approach would benefit them.

Findings from responses to item 6 of the questionnaire

Item 6 (*I am pleased with my improvement on the speaking competence after attending the course*) was designed to examine students' content with their improvement in the speaking competence after attending the course.

The pattern of the responses proves that the students in the analytic group pleased with their improvement in speaking competence after attending the course. nine students (17.65%) responded that they *definitely* pleased with their improvement. Almost half of the class, 23 students (45.10%) chose *very probably* and 13 students (25.49%) chose *probably*.

V. CONCLUSION

The research provides an insight into the interconnection between the analytic approach and students' improvement in speaking learning. This insight also helps teachers have a clear-cut decision on the implementation of the analytic scoring scheme for assessing their students' speaking performance. Moreover, the research shows that teaching and measuring students' speaking performance should be a process rather than a product. The research is a reaction to the current preferable holistic scoring approach in teaching and assessing speaking skill. The holistic scoring approach can bring some advantages to teachers in teaching and assessing students' speaking performance but introduce students some disadvantages to their autonomous process of learning speaking skill.

REFERENCES

- [1] Bachman, L. F. & Palmer, A.S. (1996). Language Testing In Practice (second edition). Oxford: Oxford University Press.
- [2] Downing, S.M. & Haladyna, T.M. (2006). Handbook of Test Development. New Jersey: Lawrence Erlbaum Associates, Inc.
- [3] Futcher, G. (2009). Rating Scales and the Halo Effect. Retrieved on July 17 2011 from http://languagetesting.info/gf/glennfulcher.php.
- [4] Hamp-Lyons, L. (1989). Second language writing: Assessment issues. In Kroll, B. (Ed.), *Second language writing* (pp. 69-87). Cambridge: Cambridge University Press.
- [5] Hughes, A. (1989/2003). Testing for Language Teachers. Cambridge: Cambridge University Press.
- [6] Hughey, J.B., Wormuth, D.R., Hartfiel, V.F. & Jacobs, H.L. (1983). Teaching ESL composition: Principles and techniques. Rowley, MA: Newbury House.
- [7] McNamara, T. (1996). Measuring second language performance. London: Longman.
- [8] Moskal, B.M. (2000). "Scoring rubrics: What, when and how?" *Practical Assessment, Research & Evaluation*, 7 (3). Retried on July 11 2011 from http://pareonline.net/getvn.asp?v=7&n=3.
- [9] Nakamura, Y. (2004). A Rasch-based analysis of an in-house English placement test. Retrieved on July 2 2011 from: http://jalt.org/pansig/2004/HTML/Nakamura.htm.
- [10] Park, T. (2004). An Overview of Portfolio-based Writing Assessment. Retrieved on July 14 2011 from http://www.tc.columbia.edu/academic/tesol/WJFiles/pdf/TaejoonParkForum.pdf.
- [11] Park, T. (2008). Scoring Procedures for Assessing Writing. Retrieved on July 14 2011 from: http://www.tc.columbia.edu/academic/tesol/WJFiles/pdf/Park_Forum.pdf.
- [12] Weigle, S.S. (2002). Assessing Writing. Cambridge, UK: Cambridge University Press.
- [13] Weir, C.J. (2005). Language Testing and Validation: An Evidence-based Approach. Houndgrave, Hampshire: Palgrave MacMillan.
- [14] White, E.M. (1985). Teaching and assessing writing. San Francisco: Jossey-Bass.

Luu Trong Tuan is currently an EFL teacher at Ho Chi Minh City University for Natural Resources and Environment. He received his M.TESOL from Victoria University, Australia in 2004. Besides his focus on TESOL, his recent publications such as Language Transfer is Cultural Transfer between Communities, *Social Sciences Review*, No. 11, 2004, pp. 60-63; and Principles for Scientific Translation, *Social Sciences Review*, No. 8, 2004, pp. 63-67; and Building Vietnamese Medical Terminology via Language Contact, *Australian Journal of Linguistics*, Vol. 29, No. 3, September 2009, pp. 315-336 show his interest in language contact and translation areas.