

Instant Text-Based Feedback Systems

The Development of a Text-Based Feedback System and Its Potential Use in Foreign Language Teaching

Yuichi Ono^{*1}, Manabu Ishihara², Mitsuo Yamashiro³

^{*1} University of Tsukuba / Foreign Language Center
1-1-1, Tennodai, Tsukuba, Ibaraki, 305-0033, Japan

² Oyama National College of Technology / College of Electrical and Information Engineering
771, Nakakuki, Oyama, Tochigi, 323-0806, Japan

³ Ashikaga Institute of Technology / Doctoral Course of Information Science and Manufacturing Engineering
268-1, Omae, Ashikaga, Tochigi, 326-8558, Japan

^{*}ono.yuichi.ga@u.tsukuba.ac.jp; ²ishihara@oyama-ct.ac.jp; ³yamasiro@ashitech.ac.jp

Abstract

This paper examines a new type of instant text-based feedback system that was implemented in a public speaking course for Japanese foreign language learners. We begin by examining past research concerning peer feedback during student presentations and its limitations. Study 1 specifically addresses the difficulty of analytic assessment in the practice of peer evaluation of low proficiency learners. After describing our feedback system and its actual use in a public speaking course for English as a Foreign Language (EFL) learners, Study 2 will demonstrate that text-based instant feedback is beneficial for both evaluators and the evaluated. The results of our study provide compelling evidence for the development of future foreign language learning instant feedback systems.

Keywords

Text-based Feedback; Foreign Language Teaching; Presentation Course; CaboCha

Backgrounds

For well over a decade, university classrooms have steadily shifted from a teaching to learner-centric approach (Barr & Tagg, 1995). The traditional “one-way” style of lecturing has been replaced by methods that engage students in the learning process. Strong empirical evidence suggests that active involvement in the learning process is important in two areas: for the mastery of critical thinking and problem-solving skills and to increase the likelihood of a student completing a program (Center for Faculty Excellence, University of North Carolina at Chapel Hill, 2009; Braxton, Jones, Hirschy, & Hartkey, 2008; Prince, 2004).

Learning styles and classroom activities enhance the learning process. Presenting a subject in a foreign language provides students with an opportunity to reconsider their approaches to critical thinking,

problem solving, collaborative learning, speaking, and writing. Instant feedback activities that engage the audience are one method of encouraging active learning for both presenters and their audience. By providing feedback immediately after a student completes his or her presentation, the experience is more authentic, and occurs precisely when he or she is the most receptive to criticism, coupled with the excitement of the reaction. Improving the quality of instant feedback activities can motivate students to become more involved in the learning process.

Traditionally, quantitative feedback has included techniques such as rating presentations on a five-point scale from very bad to excellent using popular prompts such as “How good was the student’s pronunciation?” The results are then shown instantly in a graph or picture projected onto a screen in the classroom. This approach usually involves sender-receiver communication devices dubbed “Audio Response Systems” or “Response Analyzers.” “Clicker” is the most popular name for these products, and from herein we would like to refer to its use as the “Clicker Approach.” The use of this type of feedback system in the classroom and its effects are addressed in the literature review that follows this section (Fies and Marshall, 2006; Mayer et al., 2009; Caldwell, 2007; Collick, 2008). Although most studies echo the claim that “interactivity” is key to improving class quality, no study has examined the effectiveness of a text-based qualitative instant feedback system.

This paper begins with a discussion of the Clicker Approach’s limitations in terms of practicality and ability to increase learner motivation; this is dealt with specifically in Study 1, where the device’s relationship to analytic and holistic assessments in peer evaluation

is addressed. After providing an outline of our system in detail, we demonstrate the superiority of text-based instant feedback in Study 2.

The Use of Clicker in the Classroom

Suzuki et al. (2007) examined the cause of Clicker’s widespread adoption throughout the United States. This study revealed the limitations of traditional lecture styles, which are listed in Table 1, and easily observed in many Japanese classrooms.

TABLE 1 FOUR SHORTCOMINGS OF TEACHER-CENTERED LECTURES

- Class sizes are often very large which affects a lecture’s style
- Memorizing each piece of information that is presented during a lecture is difficult
- Students have limited concentration and cannot retain each fact presented during a lecture
- There is no way for teachers to individually assess student comprehension during a lecture

Keller et al. (2007) conducted research at the University of Colorado and discovered that, “Although this breadth of use represents a tiny fraction of all faculty on campus (3%), it represents a significant fraction of the student body due to the high average enrollment of courses using clickers. In this semester, clickers were used by 10,011 unique students, which include 9,941 undergraduates and 70 graduate students” (pp. 128-129). According to Banks (2006), the role of clickers varies. It can be connected to the teacher’s PC, a content management system such as Moodle, used with whiteboards, or incorporated into PowerPoint presentations. These possibilities have largely contributed to clickers’ widespread adoption in various educational settings. The basic concepts and merits of the Clicker Approach can be applied to foreign language classrooms as well. In a public speaking course, for example, instant feedback can provide students with an opportunity to reflect upon their performance and the task’s benefits. In a public speaking course, audience feedback allows students to contemplate both the positive and negative aspects of their presentation.

Limitations: Quantitative Approach

In this section, we begin by discussing the traditional Clicker Approach and its limitations when applied to analytic assessments in peer evaluation.

Analytic Versus Holistic Evaluation

In the field of language education, holistic and analytic scorings are primary methods of assessing free writing activities. In an analytic approach, several subscales

are used to rate characteristics of a composition separately, while a holistic approach involves a single global rating to do it. In this context, holistic scoring is more advantageous than analytic approaches due to the speed at which a writing sample can be evaluated (Davies et al., 1999). One possible disadvantage of holistic judgment, however, is that evaluators may each focus on a different aspect of the writing. On the other hand, analytic scoring requires evaluators to consider specific aspects of the writing, so that identical criteria are used to evaluate the student’s performance. A simple comparison between the two scales is made in Table 2 is below.

TABLE 2 A COMPARISON IN TERMS OF SIX QUALITIES OF TEST USEFULNESS

Quality	Holistic Scales	Analytic Scales
Reliability	Lower than analytic, but still acceptable.	Higher than holistic.
Construct Validity	Assumes that all relevant aspects of writing ability develop at the same rate and can thus be captured in a single score.	More appropriate for L2 (Second Language) writers as different aspects of writing ability develop at varying rates.
Practicality	Relatively fast and easy.	Time-consuming; expensive.
Impact	A single score may mask an uneven writing profile and lead to misleading placements.	More scales provide useful diagnostic information for placement and/or instruction; more useful for rater training.
Authenticity	White (1995) argues that reading holistically is a more natural process than reading analytically.	Raters may read holistically and adjust analytic scores to match holistic impressions.

However, the disadvantage of analytic scoring, as indicated by Davies et al. above, is that it is more time-consuming. Table 2 summarizes the advantages and limitations of both methods (Nakamura, 2004).

Arguments in Favor of Instant Feedback

Choosing a scoring method is not an easy task (Nakamura, 2004). It is natural to assume that only using either scale can result in insufficiency depending on the context of the test. Yamanishi (2007) actually proposes the blended implementation into a classroom. It is natural to assume that the students will have sufficient time for making evaluation in either way and the traditional two approaches may be incorporated in the course plan.

In our scenario of providing real time feedback when the presenter is the most receptive to reaction, either

traditional approach leaves much to be desired. First, the audience has only limited time to provide feedback, say, in a few minutes, because the task is to express real time impression or comments on each presentation. Moreover, unlike a writing assessment, students cannot go back to unclear parts or where they fail to follow in a presentation.

Answering only small numbers of question items in a clicker approach might be a helpful task in this sense, since the task is simple and less time consuming. However, we are skeptical about this alternative concerning a motivation perspective. In these situations of real time peer assessment, the audience tends to make generally holistic assessments since it is very difficult especially for less confident learners to provide each rating in a limited time, even though there are only small numbers of question items. The situation will become more serious if the instructor tries to enhance quality of feedback and increase the number of question items to cover a wider range of focus of evaluation. For example, if there are five perspectives such as pronunciation, layout, organization, visual, source, with each perspective involving four questions, the audience must answer 20 questions in approximately three to five minutes. More sophisticated assessment rubrics require questions that elicit greater detail, and subsequently take more time to complete.

Another issue for analytic assessment concerns the audience's proficiency. In the present study, since the participants are EFL learners, peer assessment's reliability unequivocally depends on the audience's language proficiency. To date, studies have examined the relationship between proficiency and the reliability of an analytic evaluation. One such example is in Ono, Ishihara and Yamashiro (2013), which suggested that the analytic assessment tends to be more holistic especially for less proficient learners. It implies the possibility that less proficient learners will not be confident on their language skills with the result that their grading average becomes holistic and higher compared with high proficient learners.

In the following section, we will describe our first study.

Study 1: Proficiency and Analytic Evaluation

In order to observe how scoring tendencies differ according to proficiency level, we carried out a pilot study implementing quantitative instant feedback with participants from two different classes and

proficiency levels.

Method

Following the work of Yamamoto (2010), who extracted questions by mining data collected from a public speaking course, we drafted 20 questions. The participants answered each on a 5-point Likert scale. The statements are shown in Table 3.

The participants were students at a national university in Japan. The high-proficiency group (Class A) included 38 students majoring in the humanities and cultures; the other group (Class B) comprised 35 students majoring in information and design. Every freshman has to take placement examination to determine the class. The examination is composed of reading and listening sections on the basis of TOEFL test. The average placement test scores for both classes were 71.0 and 41.0, respectively (Max =80).

We presented both groups ten video clips that were created by students the prior academic year (Ono, Ishihara, and Yamashiro, 2012). The videos focused on various aspects of Southeast Asian culture; the subjects of the videos are shown in Table 4.

TABLE 3 QUESTION ITEMS FOR THE SURVEY

1	I made a lot of new discoveries.
2	It was informative.
3	It contained a lot of vocabularies; thus it was convincing.
4	English is easy to listen to.
5	Visuals are easy to watch.
6	It contained proper uses of colors.
7	It contained clear messages.
8	The organization is easy to follow.
9	The title was proper.
10	It is a well-researched presentation.
11	Speaking speed is proper.
12	Volume of voice is proper.
13	The speech had a proper pauses.
14	Layout of the slide is good.
15	Figures and tables are used effectively.
16	The size of the letters in the slide is proper.
17	Uses of colors are proper.
18	Story and slides are consistent.
19	Description of quotation is correct
20	Quoted information is reliable.

TABLE 4 TEN TOPICS FOR THE STUDY

- Technical Experts
- Fashion in Thailand
- My home in Thailand
- Japanese NGO and Thai flood
- Military in Thailand
- Siriraj Medical Museum
- Second life in Thailand
- Symbol of Thailand
- Thai economy

Results

The average score for each item was calculated and the results were subjected to a t-test to determine whether a significant difference existed between either group. The descriptive statistics and the results of the t-test are shown in Table 5, and reveal that the most significant items are “visual” and “aural” in nature, and that for nearly all significant items lower-class participants scored higher. These students paid less attention to the presentation’s visual or aural features rather than its linguistic and organizational elements. We believe that this is attributable to a comprehension strategy that depends on visual aids to decipher the film’s spoken or written linguistic information. The results revealed that scoring tendencies can differ between two distinct classes and proficiencies.

TABLE 5 THE RESULT OF T-TEST

	High-level		Low-level		t-value (Two-tailed)
	M	SD	M	SD	
1 I made a lot of new discoveries.	4.46	.69	4.35	.85	.055
2 It was informative.	4.36	.67	4.26	.72	.042 +
3 It contained a lot of vocabularies; thus it was convincing.	4.24	.70	4.15	2.28	.453
4 English is easy to listen to.	3.95	.90	4.04	.87	.168
5 Visuals are easy to watch.	4.18	.76	4.38	.73	.000 ***
6 It contained proper uses of colors.	4.00	.74	4.41	2.78	.003 **
7 It contained clear messages.	4.27	.71	4.27	.76	.943
8 The organization is easy to follow.	4.21	.73	4.12	.72	.111
9 The title was proper.	4.57	.65	4.49	.66	.108
10 It is a well-researched presentation.	4.44	.65	4.46	.66	.649
11 Speaking speed is proper.	4.19	.79	4.30	.74	.061
12 Volume of voice is proper.	4.03	.80	4.28	.82	.000 ***
13 The speech had a proper pauses.	4.13	.83	4.14	.81	.881
14 Layout of the slide is good.	4.09	.70	4.37	.73	.000 ***
15 Figures and tables are used effectively.	4.15	.72	4.28	.83	.026 +
16 The size of the letters in the slide is proper.	3.84	.76	4.19	.77	.000 ***
17 Uses of colors are proper.	3.96	.70	4.22	.78	.000 ***
18 Story and slides are consistent.	4.39	.68	4.49	.66	.052
19 Description of quotation is correct	3.86	.72	3.94	.85	.144
20 Quoted information is reliable.	3.95	.68	4.03	.77	.113

Next, we conducted factor analysis to determine the structural difference of the inter-relationship among questions of both classes. All the items were subject to principal factor analysis with Promax rotation. The results are displayed in Table 6. The high-level class factors are named “Content and organization,” “Visuals,” “Pronunciation and English,” and “Correctness of Information source.” Factor 1 of the low-level class, in contrast, includes Factors 1 and 2 of the high-level class. This indicates that low-level learners tend to integrate analytic standards into one abstract evaluation. This is not desirable in terms of reliability, although it may be attributable to the participants’ low proficiency or lack of confidence in determining the scores. This leads us to suggest that heavily analytic quantitative feedback could be inappropriate, especially for low-level learners.

These results clarified the limitations of analytic

quantitative evaluation for instant feedback activities in public speaking courses. The key finding was that students with low proficiency tend to evaluate items too positively and holistically for some points.

TABLE 6 RESULT OF OUR FACTOR ANALYSIS

	Class A (High)		
	Factors		
	1	2	3
17 Uses of colors are proper.	.888	-.170	-.061
14 Layout of the slide is good.	.853	-.124	-.097
16 The size of the letters in the slide is proper.	.741	.028	.000
15 Figures and tables are used effectively.	.664	-.074	.076
5 Visuals are easy to watch.	.654	.153	-.148
18 Story and slides are consistent.	.616	.127	-.004
8 The organization is easy to follow.	.548	.063	.114
7 It contained clear messages.	.508	.192	.068
2 It was informative.	.491	.126	.104
9 The title was proper.	.359	.078	.148
10 It is a well-researched presentation.	.343	.295	.153
13 The speech had a proper pauses.	.057	.834	-.110
12 Volume of voice is proper.	-.092	.804	.027
11 Speaking speed is proper.	.063	.758	-.049
4 English is easy to listen to.	-.085	.755	.058
20 Quoted information is reliable.	.011	-.054	.893
19 Description of quotation is correct.	-.029	.018	.863

KMO=.913, Total Accumulative Variance = 60.1%

	Class B (Low)			
	Factors			
	1	2	3	4
2 It was informative.	.812	-.067	-.089	.069
3 It contained a lot of vocabularies; thus it was convincing.	.761	-.061	.008	.101
7 It contained clear messages.	.727	.064	-.034	-.085
1 I made a lot of new discoveries.	.707	-.101	.033	.013
18 Story and slides are consistent.	.699	.096	-.043	-.020
9 The title was proper.	.663	-.058	.103	-.005
10 It is a well-researched presentation.	.639	.104	.039	.067
8 The organization is easy to follow.	.566	.150	.135	-.183
17 Uses of colors are proper.	-.069	.904	-.068	.020
6 It contained proper uses of colors.	-.058	.823	-.024	-.013
14 Layout of the slide is good.	.054	.671	-.057	.016
16 The size of the letters in the slide is proper.	-.083	.613	.202	-.042
5 Visual is easy to watch.	.066	.611	.117	-.018
15 Figures and tables are used effectively.	.253	.506	-.141	.122
13 It contained properly-paused speech.	-.110	.016	.850	.083
12 Volume of voice is proper.	.064	-.085	.704	.004
11 Speaking speed is proper.	.029	.057	.700	.026
4 English is easy to listen to.	.261	.024	.454	-.063
19 Description of quotation is correct.	.014	.015	.005	.894
20 Quoted information is reliable.	-.002	.010	.072	.767

KMO=.907, Total Accumulative Variance = 63.2%

System Description

The results of Study 1 compelled us to design and construct “text-data” based instant feedback systems that allowed participants to write freely about their impression of a presentation. Our expectation was that, first, low-level students would feel comfortable commenting on their points, since there was no need to pay attention to each one in the feedback activity. Secondly, we anticipated that presenters would express their true feelings regarding their performance more clearly and directly.

The system is based on natural language processing technology and basic levels of text mining. A rough overview of our system is provided in Figure 1. The system was installed in our Learning Management System as a Moodle module. This enabled participants

both outside and inside the classroom to provide feedback. Participants wrote comments freely on the LMS that were then collected by the system, which calculated frequencies categorically and produced a graph containing the results.

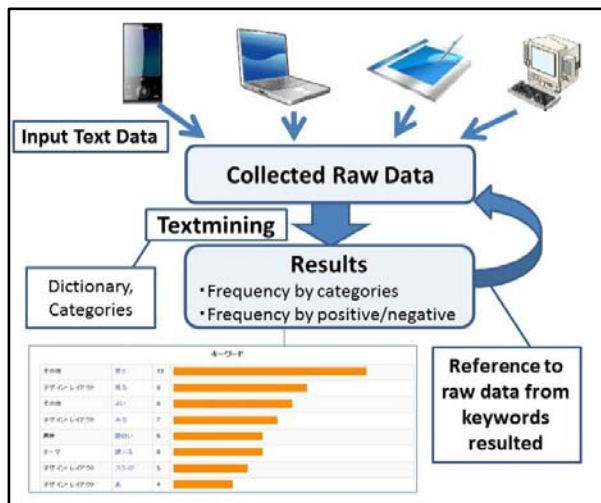


FIG. 1 OUTLINE OF THE SYSTEM

Japanese Natural Language Processing

The process of mining Japanese text usually comprises the four following stages:

- Analysis of morphology
- Analysis of structure
- Analysis of meaning
- Analysis of discourse

The first stage is unique to the Japanese language, because unlike the English language, there are no spaces between words. The analyzer first divides the text according to Japanese morphological rules. We selected an open source program MeCab for this task due to its extreme versatility.

At the next stage, the analyzer identifies dependency relations between words. CaboCha is a high-level free Japanese syntactic parser, which enables us to analyze the syntactic dependency of Japanese sentences. For example, let us consider a case concerning a relative clause.

太郎は [花子が 読んでいる] 本を 次郎に 渡した
 Taro-SUB Hanako-SUB read-ing book-ACC Jiro-DAT give-PAST
 Taro gave the book [Hanako was reading]

FIG. 2 ANALYSIS OF A JAPANESE RELATIVE CLAUSE

CaboCha analyzed the collected sentences and the results are shown in Fig. 3, cited from <http://code.google.com/p/cabocho/>.

In order to process the textual data according to our

needs, the authors created a dictionary for evaluation. For the selection of lexical items, the data collected from the pilot study was used. 50 students participated and made open-ended comments for each of the ten presentations. After analyzing these data with CaboCha, approximately 410 lexical items were selected, and two kinds of properties were assigned to each item. The first property concerned students' impressions of the presentation; the other addressed semantic categories such as "design/layout," "interest," "English," "pronunciation," "citation," and "others." A screenshot the evaluation dictionary is shown in Figure 4.

```
% cabocho -f1
太郎は花子が読んでいる本を次郎に渡した
* 0 5D 0/1 1.062087
太郎 名詞,固有名詞,人名,名,**,太郎,タロウ,タロー
は 助詞,係助詞,**,*は,ハ,ワ
* 1 2D 0/1 1.821210
花子 名詞,固有名詞,人名,名,**,花子,ハナコ,ハナコ
が 助詞,格助詞,一般,**,*が,ガ
* 2 3D 0/2 0.000000
読ん 動詞,自立,**,五段・マ行,連用タ接続,読む,ヨム,ヨン
で 助詞,接続助詞,**,*で,デ
いる 動詞,非自立,**,一段,基本形,いる,イル,イル
* 3 5D 0/1 0.000000
本 名詞,一般,**,*本,ホン,ホン
を 助詞,格助詞,一般,**,*を,ヲ,ヲ
* 4 5D 1/2 0.000000
次 名詞,一般,**,*次,ツギ,ツギ
郎 名詞,一般,**,*郎,ロウ,ロー
に 助詞,格助詞,一般,**,*に,ニ,ニ
* 5 -1D 0/1 0.000000
渡し 動詞,自立,**,五段・サ行,連用形,渡す,ワタシ,ワタシ
た 助動詞,**,*特殊・タ,基本形,た,タタ
EOS
```

FIG. 3 OUTPUT FROM CABOCHA

	A	B	C	D
37	英語・発音	聞き覚え	0	
38	英語・発音	語	0	
39	英語・発音	聞き取る	1	
40	英語・発音	発音	1	
41	英語・発音	英語	1	
42	英語・発音	聞く	1	
43	英語・発音	スピード	1	
44	英語・発音	声	1	
45	英語・発音	話す	1	
46	英語・発音	大きい	1	
47	英語・発音	話し方	1	
48	英語・発音	間	1	
49	英語・発音	迫える	1	
50	英語・発音	音声	1	
51	英語・発音	流暢	1	
52	英語・発音	丁寧	1	
53	興味	跑ぎる	-1	
54	興味	わむたい	-1	
55	興味	自分	0	
56	興味	作品	0	
57	興味	引く	1	
58	興味	良い	1	
59	興味	楽しい	1	
60	興味	面白い	1	
61	興味	いい	1	
62	興味	集中	1	
63	興味	おもしろい	1	

FIG. 4 DICTIONARY FOR EVALUATION

Roughly three minutes after the students finished writing their comments the instructor shared the results shown in Figure 5, and the instant text mining results in Figure 6. Figure's 5 include the ratio of positive/ negative words appearances, and the frequency of each semantic category. Additionally, the presenter can refer to the original text by clicking on key words as shown in Figure 6.

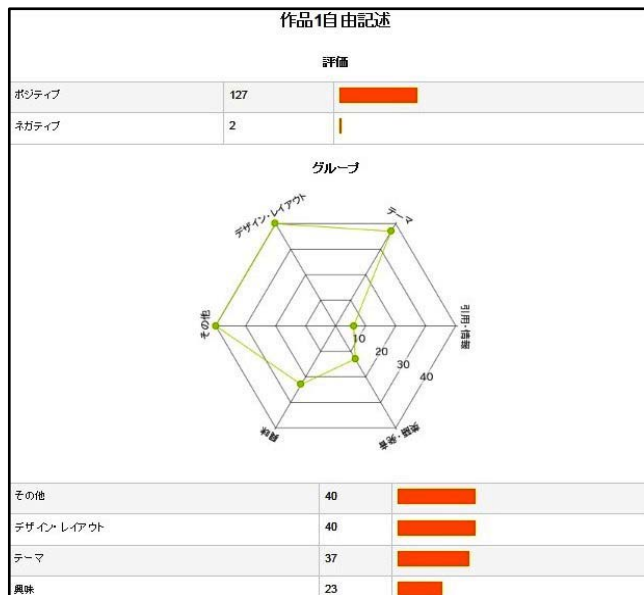


FIG. 5 OUTPUT GRAPHS

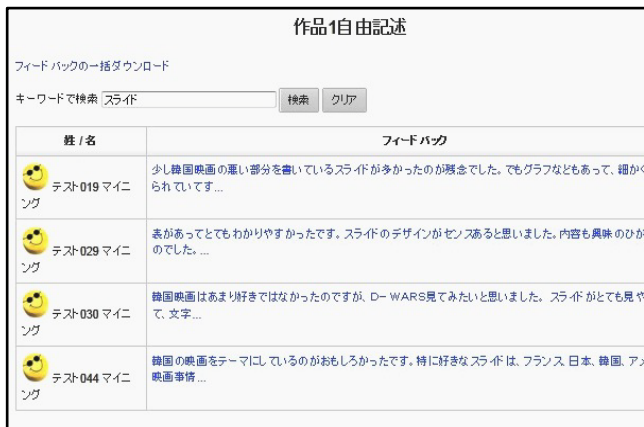


FIG. 6 FEEDBACK COMMENTS REFERENCE

Study 2: Comparison Between Two Systems

In order to gauge the students' feelings about the text-based instant feedback system, we conducted a survey concerning their impressions of the systems. Study 2 attempts to answer two research questions: (i) Is there any difference in usability? and (ii) Is there any effect on student's self-awareness or motivation?

Method

Thirty-five students from Class A participated in this study. The participants gave presentations in the class

as presenters while they acted as peer evaluators in either method. Before the introduction of text-based instant feedback system, the students were required to evaluate the 20 question items just as we did in Study 1.

The survey was conducted after the feedback sessions. After they answered the general impression of the text-based instant feedback system, they answered ten questions concerning the difficulty or degree of satisfaction they felt as an evaluator, and how their feedback might have motivated the presenter. We compared the results for both instant feedback systems. The participants also provided open-ended feedback regarding the new system. The questions are shown in Table 7.

TABLE 7 THE DIFFERENCE BETWEEN THE TWO SYSTEM

	20 items		textmining		t-value (two-tailed)	Significance
	Mean	SD	Mean	SD		
1 It was difficult to give evaluations to others.	3.167	1.090	2.958	1.197	.707	
2 It took long time to give evaluations.	3.542	1.141	2.958	1.197	1.664	
3 To was troublesome to give evaluations.	2.583	1.060	2.667	0.816	-.358	
4 Giving evaluations was beneficial for my study.	3.208	1.103	3.750	1.032	-2.184	*
5 I was satisfied with my evaluation to others.	3.583	0.974	4.000	0.933	-2.460	*
6 Their evaluations were helpful.	3.792	1.103	3.792	1.103	.000	
7 I realized how the audience feel to my presentation.	3.917	1.283	3.958	1.160	-.143	
8 The content of their feedback was convincing.	3.792	1.021	3.833	1.129	-.225	
9 Their feedback gave me a hint for the next opportunity.	3.792	0.932	4.167	1.129	2.099	*
10 I got confident at their feedback.	3.667	1.049	3.542	0.977	1.000	

***: p<.001, **: p<.01, *: p<.05, +: p<.0.1

Results

The result of the first question on the general impression of the new system was in agreement with our expectations during the system's construction: most (67 percent) of the participants responded positively (Very Good or Good) to the system. Figure 7 illustrates this.

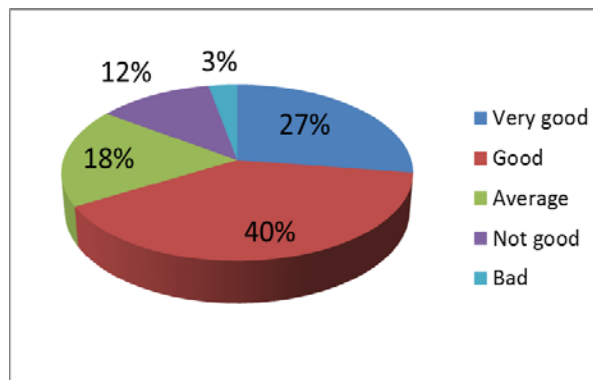


FIG. 7 GENERAL IMPRESSION OF THE SYSTEM

The statistical and t-test results are indicated in Table 7. Of them, three items were statistically significant: Item 4 “The evaluation was beneficial to my studies,” item 5 “I was satisfied with my evaluation of others,” and item 9 “Their feedback gave me good advice for the future.” This result was expected since participants were required to reflect deeply about each presentation, and express their opinions concretely; subsequently, this encouraged them to refine their critical thinking skills. On the other hand, they did not perceive a significant difference between the two feedback systems nor the difficulty performing evaluations. No significance concerning their self-confidence or anxiety was evident, either.

Next, we conducted factor analysis to observe the structural difference among question items between the two systems. All the items were subject to Principal factor analysis with Promax rotation. Items with loadings higher than 0.35 on any factor without high cross loadings were retained in the list. The result is given in Table 8. In the case of instant text-based feedback, the factor analysis resulted in two factors, while one in the case of 20 multiple-choice questions. This means that the students consider the two systems are different.

TABLE 8 RESULTS OF FACTOR ANALYSIS

Instant Text-based Feedback	Factors	
	1	2
Giving evaluations was beneficial for my study.	.915	-.013
I was satisfied with my evaluation to others.	.825	.174
The content of their feedback was convincing.	.760	.010
I got confident at their feedback.	.754	-.043
Their feedback gave me a hint for the next opportunity.	.740	.367
I realized how the audience feel to my presentation.	.688	-.176
Their evaluations were helpful.	.675	.338
It took long time to give evaluations.	.102	.784
To was troublesome to give evaluations.	-.183	.783
It was difficult to give evaluations to others.	.156	.646
accumulative variance = 61.7%		

20 multiple-choice questions	Factors
	1
Their evaluations were helpful.	.853
The content of their feedback was convincing.	.816
I realized how the audience feel to my presentation.	.808
I got confident at their feedback.	.779
It was difficult to give evaluations to others.	-.381
accumulative variance = 56.0%	

The result of our factor analysis showed that they are feeling that the two systems are different. As to the text-based feedback, two factors, say, “Satisfaction” and “Difficulty”, were abstracted in our factor analysis, while only one, say, “General merits”, was abstracted in the 20 multiple question feedback. The fact that two

factors were abstracted in the former has an interesting implication; the students are considering “Difficulty” and “Satisfaction” are independent factors of each other. The difficulty here implies that the students have to understand sufficiently enough to make some comments to the presentation, compared with the situation where the students give grades in a clicker approach. These things can be summarized as “They are difficult tasks but they are very helpful to their study.” Of course, answering 20 questions real-time is also a difficult task in a sense. The findings from our factor analysis, however, has an interesting implication that the quality of difficulty is different between the two feedbacks and that the students are satisfied with the task.

As for the results of open-ended of feedback in our questionnaire research, we were able to collect various kinds of textual data, which were then analyzed according to four points. These points were: (i) a much simpler interface compared to traditional quantitative feedback; (ii) deeper understanding of the presentation’s positive and negative aspects; (iii) usefulness of keyword reference and graphic output; and (iv) greater excitement in the class.

Lastly, we would like to show some of the open-ended comments from the participants below:

TABLE 9 OPEN-ENDED FEEDBACK FROM THE PARTICIPANTS

- The analyzing function is very exciting. The questions were reduced and simplified, which helped me concentrate more on the presentation.
- Graphs displaying the positive/negative ratios and the list of keywords were really helpful. It also helped me a lot to understand my presentation by considering the audience’s comments.
- It is enjoyable to see the evaluation and the comments from the audience instantly and the class is really exciting. I think it is important to know how the audience perceived my presentation in an objective, not subjective, way.
- In my presentation, I personally made a great effort to create the slides and layout. So, I was really pleased and encouraged that some members of the audience made comments about my slides.
- It is important to know what points to evaluate, such as content and pronunciation. Receiving comments in detail on these points was really helpful.

These comments indicate that the participants found that giving scores to 20 questions just after the presentation was a tough task and that the situation has improved because the procedure has been simplified by the use of qualitative feedback system. On the part of the presenter, the participants seem to respond positively to the comments, especially when

addressing the effort they devoted to the presentation, as illustrated in the fourth comment. They felt that their efforts were truly and instantly understood and that peers sympathized with them, which led them to feel a sense of achievement. This clarifies the advantage of the text-based instant feedback as we proposed in this paper. In summary, the practical and motivational limitations of the Clicker Approach have been solved by implementing a new system.

Concluding Remarks

In the first study, this paper examines the limitations of the traditional clicker-based approach to instant feedback activities in public speaking courses. On the basis of our questionnaire research, our second studies demonstrated that text-based instant feedback provided students with opportunities to have a deeper reflection on their own presentations and evaluate peer's presentations more deeply. We conclude that the text-based instant feedback system is advantageous to the traditional clicker-based analytic approach for foreign language learners in terms of learner's attitude and motivation.

REFERENCES

- A, Davies, A. Brown, et al. (Eds.). Dictionary of language testing. Cambridge: Cambridge University Press, 1999.
- C. Fies, and J. Marshall, "Classroom Response Systems: A Review of the Literature," *Journal of Science Education and Technology*, 15, 2006, 101-109.
- C. Keller, N. Finkelstein, K. Perkins, S. Pollock, C. Turpen, and M. Dubson. "Research-based Practices for Effective Clicker Use," 2007 Physics Education Research Conference AIP Conference Proceedings, 2007, 128-131.
- Center for Faculty Excellence, University of North Carolina at Chapel Hill. "Classroom Activities for Active Learning," Accessed August 3, 2013. University of North Carolina at Chapel Hill, 2009, <http://cfe.unc.edu/publications/fyc2.html>.
- D. A. Banks, "Audience Response Systems in Higher Education: Applications and Cases," 2006, Information Science Publishing.
- H. Suzuki, M. Takesada, T. Hikihara, K. Yamada, T. Hosokawa and A. Onodera, A. "Active Learning in the Classroom using the Response System Clicker," Report of a Physics Class in Hokkaido University in 2007, 2007, 1-17.
- H. Yamanishi. "How are High School Students' Free Compositions Evaluated by teachers and Teacher Candidates?: A Comparative Analysis between Analytic and Holistic Rating Scales," *JALT Journal*, 26(2), November 2004, 189-205.
- J. Collick, "Wireless Response Systems, Interactivity and Collaboration with Interactive Whiteboards: Towards a New Class Dynamic," *Proceedings of the Second International Wireless Ready Symposium*, 2008, 21-25.
- J. E. Caldwell, "Clickers in the Large Classroom: Current Research and Best-Practice Tips," *CBE – Life Sciences Education*, 6, 2007, 9-20.
- J. M. Braxton, W. A. Jones, A. S. Hirschy, and H. V. Hartley. "The role of active learning in college persistence." *New Directions for Teaching and Learning*, 115, 2008, 71-83.
- M, Prince. "Does active learning work? A review of the research." *Journal of Engineering Education*, 93(3), 2004, 223-231.
- R. Mayer, A. Stull, K. DeLeeuw, K. Almeroth, B. Bimber, et al., "Clickers in College Classrooms: Fostering Learning with Questioning Methods in Large Lecture Classes," *Contemporary Educational Psychology*, 34, 2009, 51-57.
- Y, Nakamura. "A Comparison of Holistic and Analytic Scoring Methods in the Assessment of Writing," *Proceedings of the Interface Between Interlanguage, Pragmatics and Assessment: Proceedings of the 3rd Annual JALT Pan-SIG Conference*, 2004, 45-52.
- Y. Ono, M. Ishihara, and M. Yamashiro, "Technology Enhanced Movie Presentation with Focus on Foreign Language Anxiety and PBL Skills," *Proceedings of the 20th International Conference on Computers in Education*, 2013, 584-588.
- Y. Ono, M. Ishihara, and M. Yamashiro, "The Construction of Instant Feedback System with the use of Text-mining in English Teaching Settings," *Research Repots Ashikaga Institute of Technology*, 47, 2013, 20-27.
- Y. Yamamoto, "A proposal of presentation evaluation skills using text mining," Master's Thesis, Nansann University, 2010.