Enhancing Critical Assessment Ability Through Oral Presentation

IWAI, Chiaki
Hiroshima City University

Reference Data:
Iwai, C. (2012). Enhancing critical assessment ability through oral presentation. In A. Stewart & N. Sonda (Eds.), *JALT2011 Conference Proceedings*. Tokyo: JALT.

The necessity of oral presentation (OP) skills for second language (L2) learners is often emphasized (e.g., Reinhart, 2002); however, it is by no means clear whether teaching such skills in class actually contributes to enhancing L2 competence itself beyond superficial improvement in oral performance, and if it contributes to it, why it does so. The present study discusses potential roles of teaching OP skills for L2 acquisition or development on the basis of empirical data collected from 24 Japanese college English learners at intermediate and high-intermediate levels. These learners were exposed to a one-semester course to train them in OP skills, and the comparison between pre- and post-training data of evaluating others' oral performances (speeches) apparently resulted in improved ability to critically assess performance, particularly at the macro-linguistic level. Based on this main finding, this study discusses the need for more such empirical studies to justify teaching OP skills not only to improve L2 learners' performance but also to facilitate L2 acquisition and development.

第二言語学習者のオーラル・プレゼンテーション（OP）スキルの重要性はよく主張されている（Reinhart, 2002）が、この技能を高めるための指導が、単なるパフォーマンスの上達だけではなく、言語の習得や発達に役立つのかどうか、もし役立つとすれば、なぜそうなのかは、必ずしも明らかではない。本研究は、中級レベルの日本人大学生英語学習者24名から収集した実証データを基に、OP技能を教えることの言語教育上の意義を考察するものである。参加学習者はOP技能訓練の授業を一学期間にわたって受講し、学期の始めと終わりに他者によって行われたスピーチの評価を行った。データ分析の結果、学習者は他者によるスピーチのマクロ言語的な問題点などに批判的評価を行う能力が高まることが確認された。この結果に基づいて、本稿では、OP技能を授業で実践することを正当化するためには、パフォーマンスの向上のみではなく、言語習得や発達の見地からさらなる実証的根拠を示す必要があることを論ずる。

The importance of oral presentation (OP) skills in English is often stressed (e.g., Reinhart, 2002), and various attempts to teach OP skills are being made in the Japanese context of teaching English as a foreign language (EFL), reflecting, for example, the Action Plan "Japanese with English Abilities" by the Ministry of Education, Culture, Sports, Science and Technology in 2003 (e.g., Fujita, Yamagata, &Takenaka, 2009). Despite such wide recognition of the importance of OPs, it does not appear, as described in the next literature review section, that oral presentations as a research topic have attracted very much attention from researchers of oral communication in second language acquisition (SLA) even though there have been various studies conducted from perspectives other than SLA. Responding to the lack of popularity of OPs as a research theme, the present study tries to achieve two objectives. The first objective is to assess a one-semester course where the author of this study taught OP

skills. The second objective, which is approached indirectly, is to discuss the importance and necessity of conducting more empirical studies on OPs from SLA perspectives.

Literature Review

Prior to the present study, academic references on L2 OPs were collected by means of several widely-used online databases such as the Educational Resources Information Center (ERIC) and Citation Information by the National Institute of Informatics (CiNii), using key words such as "oral/academic/speech presentation". The results were surprisingly sparse: no article discussing the direct relationship between teaching OP skills in L2 and its acquisition was found. Instead, several studies dealing with issues other than L2 acquisition were found. These studies are tentatively categorized here into three types.

The first type examines the psychological influence of teaching OP skills to L2 learners. In line with this research direction, Fujita, et al. (2009) reported that their trial of providing a course as a requirement for first-year university students contributed to raising learner motivation and promoting independent learning. Neff (2007) examined how English learner anxiety toward giving presentations would affect their oral performance and motivation. The second concerns how teaching OP skills expedites L2 learners' strategy use and learning autonomy. An example is Makino (2003) who assessed her Public Speaking class and concluded that learner autonomy was promoted through the goals of the class, as that was where learner awareness was directed. The third type investigates the course or syllabus design for teaching OP skills. The style and formality of OPs differ considerably, ranging from a short speech in class to an academic presentation in a specific disciplinary area, and in this group researchers and language teachers discuss course syllabi or develop textbooks for teaching OP skills (e.g., Hill & Story, 2000; Reinhart, 2002).

Needless to say, the value of these studies is undeniable; however, more direct evidence for L2 acquisition and/or development would be persuasive to rationalize teaching OP skills in class. Keeping the lack of such evidence in mind, the present study investigates the effects of teaching OP skills in class with respect to the learners' ability to critically assess speech performance by other learners. Learner ability to direct their attention to specific areas of L2 performance is related to metacognitive strategies, whose importance has been frequently discussed by researchers of language learning strategies or communication strategies (e.g., White, Schramm, & Chamot, 2007), and many strategy researchers have argued for the effectiveness of strategy training (e.g., Cohen & Macaro, 2007). The present study was conducted in line with these strategy researchers' studies to empirically verify the assumption that L2 learners' metacognitive ability, or more precisely their ability to critically perceive and analyze others' speech performance, can be strengthened within even a rather short period of classroom instruction.

Research Questions

Among the wide range of OP formats, the present study deals with presentation of a short speech and specifically investigates participants' perceptive abilities rather than actual oral performance, although performance is as important as perceptive abilities. The following two research questions were formulated for this study:

1. Does teaching OP skills have any influence on EFL learners' ability to assess others' oral speeches and speech drafts? If so, how does it influence such abilities?
2. How does learner proficiency influence the learning of OP skills?

Participants and Method

About the Class

The class investigated in this study is Academic English Expression I, an elective 2-credit course for third-year non-English major students (though they belong to a Faculty where high English skills are required) who want to develop their productive ability in speaking and writing. The class met 15 times during the semester (Table 1), with the first two weekly class sessions used for guidance. Subtracting these two, pre- and post-training data were collected at the beginning and the end of the remaining 13 classes. The total number of classes for instruction was, therefore, 11.

Participants

About 30 students registered at the beginning of the semester, but the actual participants were 24, after removing those whose pre- or post-training data were incomplete. Of these 24, 3 were male and 21 female. Gender is not a consideration for this study, so the gender balance was not controlled. Student proficiency level can be regarded as being intermediate to high-intermediate, judging from mean TOEIC® scores of 650.42 (SD = 99.10). Learners were later divided into two proficiency groups (N = 12 each) according to their TOEIC scores: a low proficiency group (LG) = 567.50 (SD = 43.72), and a high proficiency group (HG) = 733.33 (SD = 60.20). The difference between these two groups in TOEIC scores, as tested by a t-test, was statistically significant (t = 7.721, df = 22, p < .001).

Class Activities

During the 11 week training sessions (week 4 to week 14), the participants engaged in the class activities summarized in Table 1.

Table 1. Summary of Class Activities

Classes	Class activities	Instruction point
1 & 2	Guidance	—
3	Collection of pre-training data	—
4	Writing a speech draft 1	Logical structures of speeches
5	Writing a speech draft 2	Planning and topic sentences 1
6	Writing a speech draft 3	Planning and topic sentences 2
7	Writing a speech draft 4	Logical flow and connecting words
8	Writing a speech draft 5	Different types of speech drafts 1
9	Writing a speech draft 6	Different types of speech drafts 2
10	Revision of selected speech draft	Self-correction, refinement of content/form
11	Presentation practice 1	Voice, eye contact, delivery, etc.
12	Presentation practice 2	Voice, eye contact, delivery, etc.
13	Presentation and evaluation 1	Presentation by each student and peer-evaluation
14	Presentation and evaluation 2	Presentation by each student and peer-evaluation
15	Collection of post-training data	

These 11 weeks comprised three stages. The first stage (weeks 4 to 10) was devoted to writing speech drafts. Every week, one speech topic was given to the learners and they engaged in an impromptu writing activity, where they had 2 to 5 minutes of planning (the time was gradually reduced as the classes

proceeded) and handwrote a draft in 10 minutes (handwriting was chosen due to learners' slow typing speed). After the impromptu writing, explicit instruction was given regarding issues such as topic sentences, logical organization, linking words, coherence, and redundancy (about one topic per week). The textbook for the course was Hill and Storey (2000), which has sample speech videos compiled in a CD. The samples were used as a supplement for instruction. Following instruction, students participated in peer-reading activities by exchanging drafts with their classmates, and later they referred to comments from other students as a source for revisions of these drafts. After this stage, students selected the speech draft they liked best from the six drafts written for their actual performance in class.

The second stage (weeks 11 and 12) was devoted to practicing their presentation. In this activity, the learners were all paired, and they demonstrated their speeches and commented on them, changing practice partners several times and refined their drafts further as necessary.

The third stage (weeks 13 and 14) was speech presentations. All students presented their speeches with their classmates as the audience. Immediately after the speech, the audience evaluated the presentation according to six categories (speech organization, contents, clarity, eye contact, voice volume, and delivery), which were not used for this study but which influenced course grades.

Data Collection

As this study intends to examine the participants' perceptive ability toward OPs rather than actual performance, students were asked to engage in two assessment activities to examine shifts in ability. One was to watch and evaluate videotaped speeches (video), and the other was to evaluate the speech drafts (transcriptions) of these speeches (draft). The latter was included to compare the outcomes of these two assessment methods and examine whether the participants' attention would be directed more toward language aspects by giving them a chance to read a draft.

For the video assessment, two sets of three video clips (Table 2) were prepared for two data collection sessions (pre-training and post-training). Each video clip showed an oral speech of about one to two minutes, and the topics of these speeches were non-technical opinion statements such as cross-cultural experience and peaceful coexistence with other people.

Table 2. Video Clips for the Pre- and Post-Training Data Collection

	Set 1			Set 2		
		Total words	Speech rate (wpm)		Total words	Speech rate (wpm)
Slow	A	139	60.0	C	198	95.5
Mid	B	238	109.0	D	239	128.9
Fast	X	158	141.5	Y	106	151.4

Of these six video clips, A, B, C, and D were performed by students who took the author's class in the previous year, and their speeches were used in the present study after obtaining their written consent. The speaking speeds of speakers A and C were slower than B and D, and the overall performance of the former two speakers as judged by peer evaluation was statistically worse than that of the latter two as judged by their peers in the previous class. Furthermore, videos X and Y were selected from the textbook for the class. The speakers of these videos are highly proficient and well-trained English presenters in Hong Kong, and their speech performance was far better than the

former students. These two video sets were counterbalanced for data collection so that half of both the HG and LG members evaluated the speeches in Set 1, and the other half those in Set 2 at the pre-training stage; and then they were reversed at the post-training stage. To eliminate any influence of the assessment order of the videos and drafts, their orders were also counterbalanced within each proficiency group.

The videos were evaluated by the participants in two different ways. First, they jotted down evaluation comments in their native language (Japanese) while watching each video clip (hereafter comment evaluation) in an open-ended format. The data collection was conducted in a CALL room where a computer and a headset were allotted to each student, so they could replay the speech videos if necessary. The total time for the comment evaluation period was limited to 10 minutes per speech. Following comment evaluation, the participants assessed each speech video on the following six points in a 5-point Likert scale on speaking speed, naturalness, speech contents, attention to the audience, pronunciation, and accuracy (the scale evaluation).

Immediately after the video assessment, the draft assessment was given. In the same way as the video assessment, the draft assessment was carried out in two formats: an open-ended comment evaluation (7 minutes) and the scale evaluation, the latter of which comprised evaluation items on contents, logical organization, accuracy, and easiness of understanding.

To summarize, the flow of the entire data collection sessions was:

- Video assessment (3 speeches): Comment evaluation (10 minutes) Scale evaluation
- Draft assessment (3 drafts): Comment evaluation (7 minutes) Scale evaluation.

Results

Results of the Scale Evaluation

Figure 1 summarizes the overall means of the scale evaluation of the video assessment and the draft assessment. Since there were six evaluation items in the video assessment, the evaluation result of each student was converted to a numerical score with a minimum of 6 (1 point x 6 items) and a maximum of 30 (5 points x 6 items). Likewise, the scores of the draft assessment (4 items) range from 4 to 20.

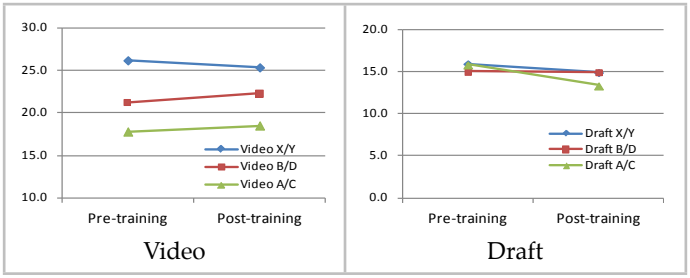

Figure 1. Scale Evaluation of Video and Draft

A two-way ANOVA (2 proficiency groups x 2 pre-post training sessions) was conducted to examine if the students' scale evaluation results changed significantly before and after the class, and no significant difference was obtained either in the video assessment or the draft assessment. These results indicate that the OP class activities did not change the learners' judgments as far as the objective evaluation in a Likert scale format is concerned. However, this was not the case in the comment evaluation.

Data Processing of the Comment Evaluation

In processing the open-ended comments by the students, first they were divided into a relevant number of pieces by idea unit (Crookes, 1990). For example, "発音はきれいだけど、話すスピードが速すぎる" (Her pronunciation was clear / but her speaking speed was too fast.) was segmented into two since this comment touches on the speaker's pronunciation and her speaking speed.

Following the segmentation of comments, two raters (the author and an experienced researcher who recently finished doctoral study) judged each comment unit on two points. One was concerned with whether the comment was given from a critical point of view (negative), from an admiring point of view (positive), or from neither (neutral). This judgment is referred to as *comment quality* hereafter. For example, the first half of the preceding example (Her pronunciation was clear) was judged positive, whereas the latter half (her speaking speed was too fast), negative. A comment like "アジアの人の英語かなと思った" (I felt her English was like Asian English) was regarded as neutral. The raters also determined which category each comment would belong to according to the taxonomies outlined in Tables 3 and 4, adapted from Reinhart (2002).

Table 3. Comment Taxonomy for Videos

Main categories	Subcategories
Pronunciation	Pronunciation/Intonation
	Voice volume
	Speaking speed
	Dysfluency (pause, hesitation, etc.)
	Others
Organization	Topic
	Logical construct
	Introduction/conclusion
	Linking words
	Coherence
	Redundancy
	Concrete examples
	Others
Appearance	Eye contact and attention to the audience
	Gesture
	Facial expressions
	Others
Language	Words and phrases
	Grammar
	Length of sentences
	Others
Other features	Easiness to understand
	Comments on speech contents
	Psychology (e.g., confidence, nervousness)
	Reliance on memo
	Native-like
	None of these

Table 4. Comment Taxonomy for Drafts

Main categories	Subcategories
Organization	Topic
	Logical construct
	Introduction/conclusion
	Linking words
	Coherence
	Redundancy
	Concrete examples
	Others
Language	Words and phrases
	Grammar
	Length of sentences
	Others
Other features	Easiness to understand
	Comments on transcript contents
	None of these

Inter-rater agreement was calculated as a percentage, and fairly high agreements were obtained as follows:

Video assessment

The judgment of comment quality: 95.11% (pre-training) and 97.07% (post-training); the judgment of comment categories: 92.82% (pre-training) and 86.00% (post-training).

Draft assessment

The judgment of comment quality: 94.27% (pre-training) and 91.63% (post-training); the judgment of comment categories: 94.24% (pre-training) and 88.04% (post-training).

Discrepancies were discussed by the two raters and final judgments made by mutual agreement.

Results of the Comment Evaluation

Comment data as explained in the preceding section were then analyzed quantitatively on the following four points: 1) the total number of evaluation comments (Figure 2); 2) ratios of comment quality by proficiency (Figure 3); 3) group means of comment types by proficiency (Figure 4), and 4) more detailed analyses of comment types (Figure 5).

1) Total Number of Evaluation Comments

The total numbers of comments increased both in video (44.1% overall) and draft (60.6% overall) conditions. The rates of increase were similar in both HG and LG. Thus, as far as the total numbers of comment counts are concerned, OP instruction induced students in both groups to comment more frequently.

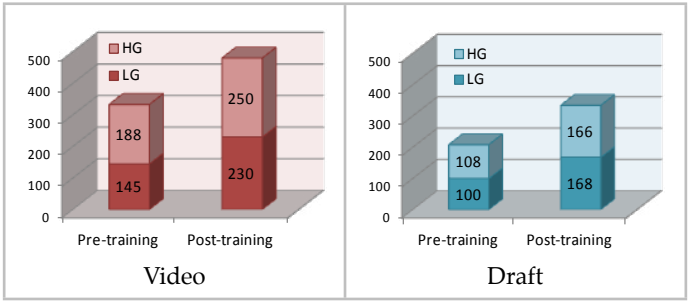

Figure 2. Total Comment Counts by Proficiency

2) Ratios of Comment Quality by Proficiency

The total numbers of comments increased, but what is more important is in what ways they increased. To clarify this, comment quality (positive, negative, and neutral) was compared between

the pre- and post-training periods across the two proficiency groups. The results show negative comments increased in video and draft, irrespective of learner proficiency. In contrast, the total counts of neutral comments dropped drastically and those of positive comments decreased slightly. The results of a two-way ANOVA confirmed these findings. Thus the learners became more critical in assessing speech presentations and drafts.

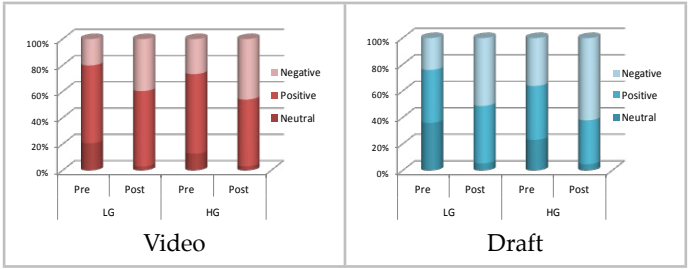

Figure 3. Ratios of Comment Quality by Proficiency

3) Group Means of Comment Types by Proficiency

Next, to scrutinize the cause of the increase in comment counts, the types of comments were examined using the taxonomies from Tables 3 and 4. Figure 4 summarizes mean total comment counts before and after the OP class per participant (comments on the three speeches and those on the three drafts were combined). As the graphs show, the category of organization underwent the largest longitudinal change in both proficiency groups, for both the video and draft conditions. Non-parametric tests revealed the difference between proficiency groups was not significant (Mann-Whitney test) and that within groups, significant pre and post differences occurred in the categories of organization and others in video (see "other features" in Table 3 for others) and that of organization and content in draft (Wilcoxon Signed Ranks test). This indicates learners' attention shifted toward macro-structural features after OP training while sensitivity to aural and visual features (pronunciation and appearance, respectively) or linguistic features (language) remained largely unchanged.

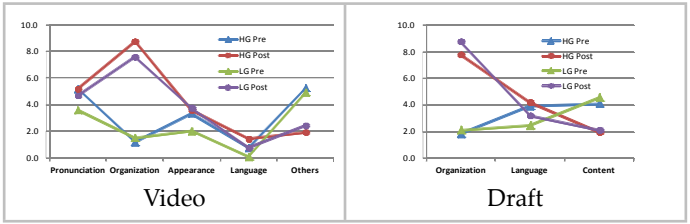

Figure 4. Mean Comment Counts in 5 Main Categories by Proficiency

4) More Detailed Analyses of Comment Types

Since organization comments in particular brought about the largest change, comments in this category were further examined according to subcategory. The proficiency groups were combined for this analysis since no significant group difference was found. Figure 5 shows that, of the 8 subcategories, the subcategory of introduction, conclusion yielded the largest change, followed by others (which included items such as asking questions of the audience effectively), logical order, examples, linking words, and coherence, with changes similar in the video and draft assessments.

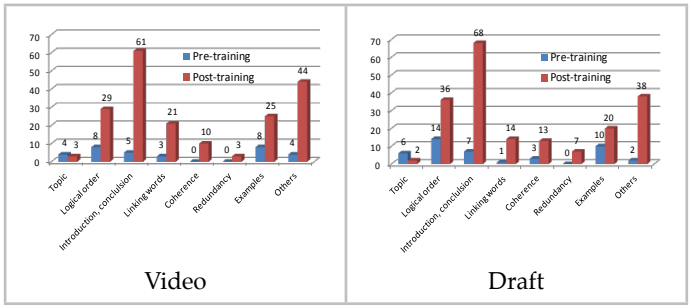

Figure 5. Comment Counts in the Organization Category

Discussion

Analysis revealed various changes in student evaluation comments after OP training although their scale evaluations did not change. Hence, the answer to the question of whether teaching OP skills has an influence on EFL learner ability to assess OPs is yes. Furthermore, the changes were salient in assessment of macro-structural features rather than sensory features (pronunciation and eye/body movements) or micro-linguistic features (grammatical, syntactic, and lexical features). The answer to the question of whether teaching OP skills differs according to learner proficiency was no as no group differences were observed, although this should be interpreted with caution since it only applies for the participants in this study.

The change in student evaluations may be superficial; a result of temporary sensitization from OP training as no delayed data was collected. The change may also be due to student interaction with the instructor after spending one semester together, which may have entered bias into the results. Acknowledging these, it is also possible that learner ability to judge what is good or bad and what is effective and ineffective in OPs was strengthened through this instruction. From an L2 acquisition viewpoint, ideally learner language ability would also have improved, but learners at or beyond intermediate level tend to improve more slowly than beginner learners. Therefore, such improvements may not be measureable over time periods as short as a semester.

This study has several potential implications, only two of which are focused on here. The first is the possibility of facilitating L2 learners' metacognitive abilities within a rather short period of time. As shown above, the OP instruction of this study was conducted for only one academic semester, and the effects on learner metacognition were statistically significant. Whether L2 language proficiency and such metacognitive abilities develop hand in hand is not clear, but they may be, to some extent, independent of one another, judging from the fact that learners with high L2 proficiency are not necessarily good oral performers, both in terms of perception and production. For such learners, the OP training as presented here could help facilitate their metacognitive abilities.

The second implication is that the findings indicate the necessity of conducting more empirical studies on the effects of teaching OP skills. While one may expect that only oral performances would improve, it is plausible that, as Yule and Tarone (1997) argued, "performance creates competence" (p. 29). If this is the case, the requisite linguistic competence for actual language use, such as pragmatic competence and discourse competence which often affect L2 use more seriously than local grammatical problems, may be enhanced by teaching OP knowledge and skills even if teaching them does not guarantee immediate improvements in performance.

Finally, it should be pointed out that the present study is exploratory, and there are several limitations that make it difficult to generalize its findings. To mention a few, only the results of quantitative analyses were presented, and the qualitative nature

of the learners' comments remains to be investigated. The number of participants involved in this research was also small, making generalization problematic. Furthermore, how metacognitive improvements influenced presentation performance was not considered here. Despite these limitations, the changes reported above are meaningful and the implications drawn from this study are worth further exploration.

Conclusion

By answering the two research questions, the present study considered the roles of teaching OP skills to EFL learners. Findings revealed there were noticeable improvements in the ways learners critically evaluate oral performances. Based on this, the study discussed the importance of developing L2 learners' ability to assess speech performance critically through OP skills teaching and the necessity of looking for more direct empirical evidence for L2 acquisition and improvement to support the value of teaching OP skills to L2 learners.

Bio Data

IWAI Chiaki, Ph.D., is a professor at Hiroshima City University. For more than 15 years, he has been examining how L2 learners use and develop strategic competence. He is also interested in the issues of interlanguage pragmatics and of English as an international lingua franca. Contact email address: <iwai@intl.hiroshima-cu.ac.jp>.

Acknowledgements

This study was supported by the academic research grant offered by the Japan Society for the Promotion of Science [Grant Code – Scientific Research (C) 22520577]. I am deeply grateful to my colleague, Carol Rinnert, for her valuable suggestions to improve the quality of this study. I also want to thank Theron Muller, the editor assigned to this paper, and the three anonymous reviewers for their insightful comments.

References

Cohen, A., & Macaro, E. (Eds.) (2007). *Language learner strategies*. Oxford: Oxford University Press.

Crookes, G. (1990). The utterance, and other basic units of second language discourse analysis. *Applied Linguistics*, *11*, 183-99.

Fujita, R., Yamagata, A., & Takenaka, F. (2009). *Gakusei-no ishikihenka-ni miru eigo presentation jyugyo-no yuyosei* [Effectiveness of an English presentation course, as seen in changing student attitudes]. *The Journal of Humanities and Natural Sciences*, *128*, 35-53.

Hill, M., & Story, A. (2000). *Speakeasy!: Oral presentation skills in English for academic and professional use*. Hong Kong: Hong Kong University Press.

Makino, Y. (2003). *Presentation-niokeru jiritsuteki gakushu-notameno gakushukankyo design* [Design of learning environment for autonomy in informative speaking]. *Japan Society of Educational Technology*, *27*(3), 325-335.

Neff, P. E. (2007). The roles of anxiety and motivation in language learner task performance. *Doshisha Studies in Language and Culture*, *10*(1), 23-42.

Reinhart, S. M. (2002). *Giving academic presentations*. Ann Arbor: The University of Michigan Press.

White, C., Schramm, K., & Chamot, A. U. (2007). Research methods in strategy research: Re-examining the toolbox. In A. D. Cohen and E. Macaro (Eds.), Language learner strategies (pp. 93-116). Oxford: Oxford University Press.

Yule, G., & Tarone, E. (1997). Investigating communication strategies in L2 reference: Pros and cons. In G. Kasper & E. Kellerman (Eds.), *Communication strategies: Psycholinguistic and sociolinguistic perspectives* (pp. 17-30). London: Longman.