(B)

2014 2015

The role of peer feedback in improving oral communication A longitudinal perspective

LEE, SHIH-CHEN NANCY

1,400,000

peer review
language-focused-instruction group   peer view group
anxiety   self evaluation   positively
  peer-review

The present study investigates the effect of peer review feedback on EFL learners' English oral proficiency development. Approximately 100 Japanese university students were divided into three groups: one control and two treatment groups. Participants in the treatment groups received either language-focused-instruction or peer-review input. After seven weeks of treatment gains in speaking proficiency were analyzed in terms of complexity, accuracy and fluency (CAF). Overall, all participants improved in CAF. Peer review group significantly improved in all CAF when compared with control group but only out performed language-focused-instruction group in accuracy. An analysis of the affective factors revealed that peer-review group had the lowest anxiety towards speaking in English and most positive self evaluation compared to other participants. The present study concluded that peer-review is effective for developing speaking proficiency of Japanese university students.

TESOL, English language education

peer review oral proficiency CAF complexity accuracy fluency

English oral communication is considered to be one of the most difficult challenges for Japanese students. In earlier studies by the researcher on the development of English academic writing ability of Japanese university students (Kaken No: 21820010 & 24720255), it was found that although the majority of participants have significantly improved their English writing ability after appropriate deployment of teaching pedagogy, many continued to struggle with basic English conversation. From participant surveys, it was found that most students had high anxiety toward speaking in English and lacked the confidence and ability to express themselves in oral communication. In addition, in interviews with selected students, it was confirmed again that Japanese university students perceive English speaking proficiency to be most difficult compared with reading, listening and writing. Therefore, it is important for practitioners and researchers to continue to investigate strategies and pedagogies to effectively develop English speaking proficiency for more Japanese learners.

Research on speaking proficiency development has a long history in the field of English education. In the EFL context, speaking proficiency development research is also gaining increasingly more attention due to the demand and English communicative ability in the increasingly globalizing society. As the result, increasingly more communicative English classes are being implemented into the university mandatory English curriculum. One response to this trend is task-based language teaching (TBLT), which suggests a shift away from the traditional direct teaching method to a new focus on task completion, which provides authentic experience and interaction for students. However, this type of communicative language teaching has been criticized for its overemphasis on authentic interaction and for overlooking the importance of paying attention to linguistic forms.

Previous studies on speaking proficiency development in the field of task-based learning and teaching have predominantly focused on the effect of planning and repetition (Bygate, 1996; 2001; Bygate & Samuda, 2005; Foster & Skehan, 1996; Ortega, 1999). The positive effects of planning and repetition on speaking production have been confirmed in previous literature (e.g., Foster & Skehan, 1996; Ortega, 1999). Besides Elder and Iwashita (2005) who found no positive effect on all three aspects of CAF, empirical results indicate that planning develops learners' fluency in most cases (e.g., Tavokoli & Skehan, 2005), develops learners'

syntactic complexity in some cases (e.g., Yuan & Ellis, 2003), but almost never develops learners' accuracy (e.g., Mochizuki & Ortega, 2008). Planning improves fluency because it improves working memory and enables learners to grasp a better conceptualization of the task beforehand (Ellis, 2005; Foster & Skehan, 1996; Ortega, 1999). This improvement can be characterized by reduction in dysfluencies and repairs (Ellis, 2009a). Thus, planning and repetition alone are insufficient to guide learners' attention to the target forms (Schmidt, 1994) for developing accuracy.

Peer feedback is increasingly becoming acknowledged for being an effective strategy as it has positive effects on both reviewers and reviewees when students provide and receive feedback (Lundstrom & Baker, 2009; Lee & Tajino, 2008; Lee, 2010). Peer corrective feedback may increase learners' awareness of linguistic forms as they are encouraged to notice non-target like utterances in their peer's production when providing comments and also to notice corrections provided by peers when receiving comments. Currently, peer review is being predominantly used in writing classrooms and improves the quality of writing as well as critical thinking, learner autonomy and social interaction between students (Hirose, 2008, 2009; Lee, 2010). Due to positive results from writing related research, it can be considered that the usage of peer review might also have positive effects on speaking development. Despite the usage of peer feedback in oral communication development is currently a relatively new concept, it has already been indicated that peer review in language learning in general enhances learner motivation, peer interaction, autonomy and self-confidence (Chu, 2013).

Besides the lack of recognition of the effect of peer review in speaking proficiency development, many speaking proficiency studies also lacked external validity because they were restricted to small size samples and cross-sectional designs due to the complex nature of collecting and analyzing spoken data. For instance, a study conducted by Mochizuki and Ortega (2008) looking at the effect of focus-on-form instruction on speaking production and focused on the changes of the relative clauses. However, they only had once off treatment. A larger number of longitudinal studies with larger sample sizes need to be conducted.

The first purpose of present research is to investigate the effectiveness of peer

feedback on speaking development in terms of fluency, accuracy and complexity growths (Ellis, 2005; Skehan, 2009). Peer review is an overlooked area in speaking proficiency development, especially in the EFL context because of the controversies surrounding students' ability to give receive adequate feedback comments. Therefore, this research aims to investigate the effectiveness of giving and receiving peer feedback comments in order to add to the body of literature on what and how to provide more effective peer review experiences for developing speaking proficiency in Japanese university level students.

The second purpose of the present research is to examine longitudinal growth of EFL students' speaking proficiency from the effects of peer review. While many cross-sectional studies of speaking proficiency development research have been conducted in the TBLT classroom, a body of longitudinal studies is still lacking.
It aims to investigate weekly trajectory changes in learners' speaking proficiency over nine weeks during the conduct of normal classroom time. Therefore, the present research would be able to contribute to the need for more longitudinal results for speaking development research and to add new perspectives to the field of TBLT.

The third purpose of the present research is to the speaking proficiency development of EFL Japanese university students by employing a larger sample size than most previous studies. Past speaking proficiency development studies often involved a limited number of participants (e.g. Larsen-Freeman, 2007, five participants), which limited the types of statistical analyses that can be employed. The present study aims to investigate speaking proficiency development of approximately 100 Japanese university students.

The final purpose of the present study is to investigate the effect of peer review on learners' affective factors especially learners' anxiety to speaking in English. The present study aims to explore the possible bilateral influence of anxiety and speaking development on each other, to contribute to a better understanding of the influence of affect on EFL learners' speaking development. In addition, it has been argued that affective factors such as speaking anxiety might be more appropriate indicators of language success than actual skill based performance of learners because of the possible influential role affective factors play in determining behaviors

(Schunk, 1991). Therefore, the fourth purpose of this study serves as an additional dimension for measuring changes in learners' English speaking proficiency.

In order to investigate the influence of peer review feedback on the development of speaking proficiency, quantitative examination was conducted where students narrated cartoon pictures in English after being exposed to 1) peer review feedback, 2) language focused instruction, or 3) no input.

The participants were approximately first-year EFL learners attending a university in Western Japan. All students at this university have high academic ability from passing standardized national Japanese center test. In addition, most students are considered to have medium to high English writing and reading proficiency levels. However, speaking proficiency among students is very diverse. Participants of this study came from five intact mandatory English speaking classes taught by the researcher. The five classes were randomly allocated into the following three groups: control, language focused instruction and peer-review. It was observed by the researcher that participants varied slightly in their motivation to learn English, and also their speaking proficiency level. According to classroom observations by the researcher, peer-review participants had slightly higher motivation than participants in the other two groups. There was no major difference in participants' English speaking proficiency from researcher's subjective observation.

Cartoon description was chosen over story narratives because cartoon is more structured, which helps to standardize content and difficulty between the subjects (Kormos & Denes, 2004; Mora, 2006). The cartoon pictures were selected from a commercial *Eeiken* Pre-Level 1 speaking test preparation textbook. For the pre-test and post-test, the same four-picture cartoon test was used. For the seven weeks of treatments, seven different four-picture cartoon tests were used, one cartoon per week.

Pilot tests were conducted to investigate the appropriate difficulty level of the cartoon pictures. Two teaching assistants were asked to evaluate the suitability of the cartoons. The four-picture cartoon (Pre-Level 1) was reported to be most appropriate as the greater number of pictures assists subjects when formulating stories. In addition, Pre-Level 1 was considered to be appropriate for higher-level university students according to the *Eiken* official website. A total of eight four-picture

cartoon pictures were chosen from a collection of 14 pictures.

Participants were asked to narrate stories in English using four-picture cartoons before (pre-test instrument) and after (post-test instrument) the 10-week treatment period. In order to familiarize subjects with test delivery and recording devices, subjects were asked to do the task in Japanese one week before the pre-test. The experiment was conducted in computer classrooms during normal class time every week. The narrations were performed and recorded individually by the participants on the computers using *Movielecom*.

Participants were given 60 seconds planning time and were asked to narrate the four-picture cartoon for a maximum of 120 seconds. They were asked to repeat the narration twice consecutively with planning before each narration. Participants were given the same four-picture cartoon for pre-test and post-test. During treatments, participants were given different cartoons once a week for seven weeks. In addition, participants in peer-review and language focused instruction groups were exposed to inputs between the repeated narrations. Participants in the language focused instruction group were exposed to direct teaching where the researcher explained one key grammar point and demonstrated key phrases. Participants in the peer-review group listened to their partner's recorded narration, gave feedback to each other before recording before recorded their second narration. Participants in the peer-review group received training on how to give and receive peer review feedback prior the beginning of the study. Participants were given the flexibility to give comments in L1 or target language or a combination of both. A peer review checklist focusing on the three constructs of speaking proficiency (complexity, accuracy and fluency) was given to the participants to assist them with formulating comments.

The total speech time (in seconds) of each narration was measured from the first to the last word. The initial silence before the first sound and the period of silence after the last sound as well as utterances not related to the cartoon such as the subjects' self dialogue or task completion statements (e.g. I am finished) were subtracted before the analysis. Dysflunecy markers such repetition, restart, rephrasing and partial words as well as filled pauses and silent pauses under 3 seconds (Riggenbach, 1991) were included in the calculation because they are considered to take parts of the natural speech even in L1.

In speaking proficiency research, the relationship between fluency, accuracy and complexity has always been controversial (Foster & Skehan, Bygate, 1998; Iwashita, McNamara, & Elder, 2008). Although it is generally assumed that proficient L2 speakers have the ability to perform oral tasks using appropriately complex morphosyntax, accurate grammar, and an appropriate level of fluency (Ellis, 2009a), these three constructs interact with each other differently depending on the speakers' proficiency level, and on the definition of these three constructs. While most researchers (e.g., Bygate, 1998; Ellis, 2005a; Foster & Skehan, 1996) have suggested trade-off effects between the three constructs, other researchers (e.g. Robinson, 2001) have argued for simultaneous improvement of the constructs. Trade-off effects occur due to limited working memory capacity. In such cases, language learners would compensate for inadequacy in one construct by performing better on the other two constructs (Ellis, 2009a). In addition, learners who prioritize one construct can perform less well on the other two (Ellis, 2005b). For instance, speakers who prioritize accuracy might produce linguistic forms they have internalized previously and be resistant to trying more complex and less familiar forms (Skehan, 2009a). On the other hand, Robinson (2001) has argued that the simultaneous improvement of all three constructs is possible when the cognitive demand of the task is increased.

Therefore, multi-dimensional analysis was used for the present study. Complexity is the measurement of the extent to which the language produced is elaborated and varied (Ellis, 2003). Despite being the most controversial construct of the three, the present study only focused on syntactic complexity. Accuracy is the ability to produce target-like and error-free language (Housen, Kuiken & Vedder, 2012) which is often measured by identifying deviations from native speaker norms in term of errors. Although the definition of accuracy is straightforward, it is controversial as to which aspects of grammar should be used to assess accuracy. Fluency is the most straight forward construct of the three and focuses on the speed of the language produced without unneeded pausing or dysfluency markers, such as hesitations, false-starts, or reformulations (Ellis, 2003).

Depending on how speaking proficiency is conceptualized, its measurements also change accordingly (Ellis, 2009). Complexity measures how advanced the language is used by learners (Skehan, 2009). Norris and Ortega (2009) argued that three aspects of complexity needs to be measured: global complexity, subclausal complexity, and complexity by

subordination. In this study, global complexity was measured by mean length of T-units, subclausal complexity was measured by mean length of clause, and complexity by subordination was measured by clauses per T-unit ratio. T-units were used rather than AS-units because it has been argued that the results produced by AS-units are unstable and difficult to generalize (Foster, Tonkyn, & Wigglesworth, 2000; Wigglesworth, 1997). In addition, T-unit based measurements have been considered to be suitable for analyzing monologue narrative tasks.

Accuracy, which in this study refers to grammatical accuracy, was measured in terms of global accuracy (Foster & Skehan, 1996) which was measured by error-free T-units. Error-free T-units are T-units free from grammatical errors, which include specific errors, word order errors, morphological errors, and the omission of obligatory forms such as pronouns.

Global accuracy measure was considered the most comprehensive for this study because all errors were considered despite the difficulty involved in establishing consistency in coding errors (Iwashita, Brown, McNamara, & O'hagan, 2008). Fluency was by speed fluency and repair fluency (Chambers, 1997; Lennon, 2000; Towell, 2002). Speed fluency refers to speech rate, which is the number of syllables produced per minute (or second). It is considered a valid quantitative fluency measure because it standardizes differing word length, so the results can be compared across studies (Kormos & Denes, 2004; Mora, 2006). Repair fluency refers to the number of self-repairs, which are indexed by a combination of self-correction measures such as reformulations, repetition, false starts, and replacements. Breakdown fluency, which is measured by silence-related indices was excluded from the present study because most pauses allocated in the narrations were under three seconds because participants were asked to speak as much as possible within the short narration time. Due to the pushed output, breakdown fluency was considered to be inappropriate for the purpose of this study. For self-repair, a lower number indicate higher fluency where a higher number in speech rate indicates higher fluency.

In addition to CAF measurements, an online survey was delivered to participants after pre test and post test using Surveymonkey to measure potential changes to participants' beliefs towards speaking in English. A Classroom Speaking Belief Questionnaire (CSBQ) which aims to measure the participants' anxiety toward speaking English in the English EFL classroom was adapted from Horwitz, Horwitz, and Cope's study (1986). It is believed that learners' anxiety is related to their language performance (Usaha & Yaikhong, 2012). Therefore, it can be considered that having an understanding of learners' anxiety toward speaking English can contribute to a better understanding of their speaking performances and development.

The 39 items in the questionnaire were grouped into three categories: Classroom Speaking Anxiety (17 items), Perceived English Speaking Self-Competence (11 items), and Desire to Speak English (11 items). The questionnaire was piloted with ten Japanese university students from the target university and was translated from English into Japanese.

To examine whether or not peer-review is effective for developing EFL university students' speaking proficiency, changes in participants' complexity (mean length of T-unit, mean length of clause, and clause per T-unit ratio), accuracy (percentage of error-free T-units), and fluency (speech rate and self-repairs) were compared with changes in learners who received language focused instruction input or no input. The gain scores of three groups were calculated by comparing results of the post-test to pre-test.

Overall, it was found that all participants improved in all CAF measurements of speaking proficiency with minor exceptions. This suggests that repetition alone contributes towards development in speaking proficiency as suggested by previous studies. However, there were significant differences between the treatment groups (peer-review and language focused instruction inputs) and control group (repetition only).

Peer-review group significantly improved in all CAF when compared with control group. For complexity, participants in the peer-review group improved considerably in the mean length of T-unit, mean length of clause as well as clauses per T-unit ratios. For accuracy, peer-review group improved significantly in the percentage of error-free T-units. For fluency, peer-review group improved slightly in speech rate but did not improve in self-review. When compared with language focused instruction group, peer-review group improved remarkably in accuracy in the percentage of error-free T-units as well as in one measure of complexity in mean length of T-unit.

Most outstanding improvement was recorded in the percentage of error-free T-units, where peer-review group improved almost

30% more than the language focused instruction group and 16% in the task repetition alone group.

For speech rate, differences between the three groups were small. The peer-review group improved the most with 1.72 syllables per second more than control group (task repetition alone) with 1.60 syllables per second improvement more than language focused instruction group.

An analysis of the affective factors revealed that participants in this study reduced in their anxiety to speak in English when compared the results from post-test to pre-test. This suggests that practice and more time on task reduces learner anxiety when speaking in the target language.

Peer-review group had the lowest anxiety towards speaking in English in terms with classroom speaking anxiety, perceived speaking self-confidence and desire to speak English outside of classroom. The biggest reduction in classroom anxiety for the peer-review group was with participants' belief towards making mistakes when talking in English with classmates as well as being nervous when speaking in pairs and in front of the classroom. With perceived English speaking self-competence, peer-review group made the biggest improvement in the perceived ability to give English presentation to small group of classmates. However, this improvement was not statistically significant when compared with control and language focused instruction groups. With the desire to speak English, peer-review group improved the most in terms with the desire for overseas travel as well as the desire to make friends with people from English-speaking countries.

To sum sup, three types of speaking conditions were compared in the present study: peer-review, language focused instruction and task repetition. Results confirmed that peer-review significantly develops speaking proficiency especially in the accuracy component of speaking. Trade-off effects might have occurred when learners prioritized accuracy, they did not improve significantly in complexity and fluency. Gains from the effect of repetition were universal to all participants in this study. Therefore, task repetition alone is beneficial for developing proficiency but to make greater grains in accuracy, peer-review inputs are desirable in the task-based language learning classrooms. In addition, peer-review is effective at reducing learner anxiety towards speaking in English and potentially increases their perceived English speaking efficacy.

These findings have an implication for Japanese tertiary English education especially for the English communicative classrooms where increasingly more task-based learning activities are being implemented. Peer-review has positive influences on students' speaking proficiency development as well as their beliefs towards speaking in the target language.

2
1. Lee, N. S. C. (forthcoming). Effects of language-focused instruction on oral proficiency development: The growth of complexity, accuracy and fluency.
2. Lee, N. S. C. (2015). My personal experience with using the CALL classroom, *Cybermedia Forum*, 16, 38-39.
5
3. Lee, N. S. C. (2015, June). Effects of explicit instruction on EFL speaking proficiency development, Faces of English Conference, Hong Kong.
4. Lee, N. S. C. (2015, Feb). Effects of explicit teaching on the speaking proficiency development of Japanese University students. Temple University Colloquium. Osaka.
5. Lee, N. S. C. (2014, Sept). Effects of explicit instruction on oral proficiency development. Kyoto JALT, Kyoto.
6. Lee, N. S. C. (2014, Aug). The effect of explicit teaching on EFL oral fluency development. AILA Congress, Brisbane, Australia.
7. Lee, N. S. C. (2014, Apr). The influence of explicit teaching on EFL speaking fluency development. 2014 International Conference on Applied Linguistics & Language Teaching. Taipei, Taiwan.

(1)

LEE, SHZH-CHEN NANCY

60512308