学 位 論 文 要 旨

Chapter 1: Introduction

When we, as English teachers, evaluate learners' English compositions, we use either a holistic evaluation method or an analytic evaluation method. One well-known holistic criterion is the TOEFL Writing Scoring Guide. The ESL Composition Profile is also well-known as an analytic criterion. However, if we evaluate novice learners' compositions using these criteria, most of the learners' scores are almost the same, and we cannot distinguish between them. Is it ideal to evaluate these compositions without distinguishing between each learner's good points? The answer is that we should evaluate them by considering every learner's strong points and weak points. Then, we Japanese English teachers should evaluate Japanese EFL learners' compositions taking into account what they want to write. In order to evaluate novice learners' compositions clearly and accurately, we need new criteria for novice EFL learners.

In this dissertation, the author reviewed the existing criteria and developed new writing criteria for novice EFL learners.

Chapter 2: Literature Review

In this chapter, the author reviewed some writing evaluation criteria used for ESL and EFL learners. She indicated their characteristics, the merit and demerit of holistic evaluation criteria and analytic criteria, and how to design an assignment rubric. Furthermore, severastudies on the reliability of some rubrics for the evaluation of Japanese EFL learner compositions were discussed.

The author also examined the features of English sentences which university students in Japan wrote. In her research, she had investigated students' verb tense and aspect errors(Baba, 2002a), cohesion (Baba, 2003), 10 morphemes (Baba, 2005b), and sentence patterns (Baba,

2006a, b). She suggested that Japanese EFL students often misuse the present progressive tense, confound passive voice and perfect tense, and sometimes omit 'be,' the main verb, subjects, or objects. She also conducted a survey on slow learners' English compositions (Baba, 2008), and the results indicated the same tendency as her previous research.

Chapter 3: Evaluating Japanese EFL Learner Compositions: TOEFL Writing Scoring Guide

In this chapter, the author identified the threshold level of novice Japanese EFL learners' English proficiency to which the TOEFL writing scoring guide can be applied. Eight Japanese English teachers evaluated 200 learners' English compositions, using the TOEFL writing scoring guide. She computed the correlation between the average scores of compositions and learners' TOEIC Bridge scores to determine the threshold level. The results indicated that the TOEFL Writing Scoring Guide can be useful to Japanese EFL learners whose TOEIC Bridge scores are greater than 130.

The author also investigated the appropriate evaluation standard for Japanese EFL learners' compositions. In order to determine this standard, eight Japanese English teachers completed the questionnaires. The results indicated that the evaluation standard should contain the following items: vocabulary, grammar, word count, content, organization, attitudes for communicating in English, word choice, and teachers' treatment of errors repeated many times in the same composition.

Chapter 4: Development of English Writing Criteria for Japanese EFL Learners:
Setting Criteria for Novice EFL Learners

The purpose of this study was to develop English writing criteria for Japanese EFL learners, and to examine the criteria reliability and validity. The study participants were 12 graduate students majoring in English education. With the new criteria developed by the author, the participants evaluated 30 English compositions written by Japanese EFL learners and answered questionnaires about the criteria. The author examined the correlation between the average scores of the learners' compositions and their TOEIC Bridge scores, and between the average scores of their compositions and word count in their compositions.

The correlation between the average scores of their composition and their TOEIC Bridge scores was .571** (p < .01). The correlation between the average scores of learners' composition and word count contained within their compositions was .825** (p < .01). These results indicate that the criteria are reliable for evaluating novice EFL learners' compositions. The questionnaire responses indicate that the grammar and content items are easy to evaluate, but those of vocabulary and organization pose some issues. Clearer explanations for using these items are needed.

Chapter 5: Evaluation Criteria for Novice Writers: Reconsidering Word Count

In this chapter, the author investigated whether word count is appropriate as one of the new criteria's categories.

The research questions of this study were the following: (1) What is the word count for novice writers for a 20 minute composition? (2) Does the word count in novice writers' compositions correlate with their English language proficiency? The participants were 147 university students in Tokyo.

In the first week, the participants took the TOEIC Bridge test. In the second and third week, they wrote English compositions with a four-colored ball-point pens for 20 minutes during the lesson. While writing, participants were asked to change the color of the ball-point pen every five minutes. Once the data collection was completed, the author counted the total number of words in each composition. The results indicated that most participants concentrated on writing in the first five minutes and they wrote more words during that portion of the 20 minutes than they did in the time remaining. They wrote what they wanted to write for the first five minutes, and then they added several sentences.

The correlations between TOEIC Bridge scores and word counts contained in the compositions was .397** (p < .01) and .380** (p < .01). Learners whose TOEIC Bridge scores were less than 130 were not independent writers and the word count varied depending on the composition theme, indicating that the word count varied based on whether they had the language or ideas. Thus, the author had to reconsider whether she should include word counts within the learners' compositions as one of the items of the criteria.

Chapter 6: Development of New Writing Evaluation Criteria for Novice EFL Students: Modifying the Criteria and Its Effects

The purpose of this study was to verify the reliability and validity of the revised writing criteria for novice EFL students. The participants for this study were 11 English teachers who were asked to evaluate 100 learners' writing samples using the new criteria. After the evaluation, the author analyzed inter-rater reliability among the 11 participant raters. She also analyzed the correlation between learners' TOEIC Bridge scores and writing scores, and between the word count contained within learners' compositions and learners' writing scores. Then, using Generalizability theory, she estimated the variance components and generalizability coefficient, and simulated the number of raters, which is necessary for reliable measurement.

Inter-rater reliability among the 11 participant raters was very high, and the generalizability coefficient was also high at 0.95. Three items were included in the criteria: vocabulary, grammar, and contents. If the number of items was greater than three, then inter-rater reliability was lower. According to the Decision study, more than three raters are needed for reliable measurement of inter-rater reliability.

The correlation between the scores and TOEIC Bridge scores was .508 (p < .01), and the correlation between the scores and the total number of words was .719 (p < .01). In order to explain the criteria, the author performed a multiple regression, and R^2 was 65.4%. Therefore, the criteria were appropriate for novice learners. In addition, learners who wrote more sentences had more grammar and spelling errors than those who wrote fewer sentences, although they made the same errors. In the manual of the criteria, the author explain how to treat such cases.

Chapter 7: Confirming the Concurrent Validity and Reliability of the New Writing Criteria for Novice EFL Learners

The purpose of this study was to confirm the concurrent validity and reliability of the new writing criteria. The participants were six graduate students majoring in English education. They were asked to evaluate 30 English compositions written by EFL novice learners using the new writing criteria and ESL composition profile (Jacobs, et al, 1981). After that, the author investigated the inter-rater reliability among five participant raters, and analyzed the correlation between the scores of the new criteria and ESL composition profile. Furthermore, using Generalizability theory, she estimated the variance components and generalizability coefficient.

Inter-rater reliability between the six raters was very high, and the Cronbach's alpha reliability was found to be .944. The reliability was expected to become stable if the number of raters increased. Furthermore, the result of the G-study showed that the English-writing abilities could be easily distinguished, thereby indicating that the new criteria are appropriate for novice learners.

Chapter 8: General Conclusion and Pedagogical Implications

In this dissertation, the author attempted to develop the new writing criteria for novice Japanese EFL learners. She examined the validity and reliability of the criteria, and found that they were appropriate for novice learners, and were useful not only for university teachers, but also secondary school teachers. It is hoped that the criteria are used by English teachers as good tools for teaching writing.