The present study explores a theoretical framework for observing lessons of English as a Foreign Language (EFL) and aims to develop a lesson observation instrument, which will be utilized to evaluate and improve EFL lessons in junior high schools in Japan. Under the recent educational reform mandate, teacher evaluation is conducted in various local school districts for the purpose of quality control. However, teachers tend to be evaluated in a top-down way and it seems that the evaluation system does not necessarily provide teachers with opportunities to reflect upon their teaching and participate in their own professional development. We strongly believe that lesson observation and evaluation should be utilized to promote teachers' own professional development and encourage their own understanding of the classroom practice. Although, the discussion will be restricted to English language teaching, we would like to design a lesson observation instrument, which can be used by teachers for their own professional development.

We will base the development of the lesson observation instrument on the one called the Sheltered Instruction Observation Protocol (SIOP) (Echevarria, Vogt and Short, 2009) The SIOP was developed and used in United States of America developed by the Center for Research on Education, Diversity and Excellence California University at Berkley (CREDE) and Center for Applied Linguistics (CAL). It has been developed in order to evaluate a Sheltered Instruction (SI), where English language learners (ELLs), whose L1s are other than English, learn the content area knowledge as well as the language. Successful SI are supposed to lead the ELLs to the main stream classes of the school. The SIOP consists of thirty assessment features, which are grouped into eight main features; they are Lesson Preparation, Building Background, Comprehensible Input, Strategies, Interaction, Practice/Application, Lesson Delivery, and Review/Assessment. The SIOP (a) presupposes high quality SI lessons, which can be validated by modern theories of teaching and learning, (b) involves teachers as well as administrators, principals, colleagues, in the process of evaluation and further planning of lessons, and (c) provides a reliable and valid observational sheet.

In the present study we would like to incorporate the framework of the SIOP into teachers’ lesson evaluation, which can be utilized in the Japanese EFL classroom by considering the contextual differences between the schools in the US and those in Japan.

To modify the SIOP and develop a lesson assessment instrument, which suits Japanese junior high school contexts, the
following steps were taken. First, the thirty items in the SIOP were translated into Japanese by the author, and the list was discussed and examined by several people. As a result, we obtained 61 assessment items under the following eight main features, (1) Lesson Preparation, (2) Building Background, (3) Comprehensible Input, (4) Strategies, (5) Interaction, (6) Practice/Application, (7) Lesson Delivery, and (8) Review/Assessment. “Team-teaching” was added because it is a major characteristic of Japanese English lessons. To examine the usefulness of the sixty-two items for assessment, we asked nineteen English language teachers to join the survey and rate the items. They were all Japanese English language teachers, all of whom had taught at junior high schools for more than 10 years. After the survey, we calculated the mean scores of usefulness and standard deviations (SD) of the rating for each assessment item. To diminish overlap among the items and to increase the feasibility of the assessment tool, the initial 61 items were reduced to thirty-two items. Among the nine features, “scaffolding” included in “Strategies” was decided to be separated because scaffolding by the teacher is considered very different from strategies. As a consequence, our assessment instrument comes to thirty-two features in ten main headings.

To test the reliability of our observation instrument, we put it to use in the Japanese junior high school context and calculated the interrater reliability, following the procedure described in Guarino, et al. (2001). Three video-taped lessons were observed and rated by two raters independently. After collecting the evaluation sheets, Pearson correlations coefficient was calculated for all three lessons. The mean of the coefficients of the three lessons was .907, which is considered an appropriate estimate of interrater reliability. However, individual coefficients of the rating of the three lessons varied from .818 for Lesson A to .943 for Lesson C. A detailed analysis suggested that the inconsistency was due to the varied judgment for interaction and scaffolding, which tended to require subjective judgment by the raters. Although some inconsistencies of scores among the raters were observed, we believe our lesson observation instrument has been sufficiently modified to the Japanese context to evaluate lessons in a relatively reliable way.

We believe that lesson observation incorporating teacher intention will enrich the classroom community and provide a means for teacher development. In conclusion, we suggested the refinement of the lesson evaluation instrument be continued and teachers need to be involved in the development process.

主任指導教員 吉田達弘