

Current practice of Classroom Speaking Assessment in Secondary Schools in South Korea



**THE UNIVERSITY
OF QUEENSLAND**
AUSTRALIA

by Sujin Lee
(41938553)

The University of Queensland

Supervisor: Dr. Noriko Iwashita

A Thesis Submitted in Partial Fulfillment of the Requirement for
the Degree of Masters of Arts (MA) in Applied Linguistics, The University of
Queensland

June, 2010

STATEMENT OF ORIGINALITY

I declare that the materials in this thesis have not been submitted previously or concurrently in the whole or as part of requirements for a degree to the University of Queensland or any other educational institutions except as fully acknowledged within the text.

I also certify that the work presented in the thesis is to the best of my knowledge and belief, original except as acknowledged in the text. Any help that I have received during the preparation and completion of this thesis has been acknowledged. In addition, I certify all information sources and literature used are indicated in the thesis.

Date: 2010. 6. 15

Signature:

Sujin Lee

ACKNOWLEDGEMENTS

I would like to express my immense gratitude to all those people who have given their invaluable support and assistance.

In particular, I am deeply indebted to my supervisor, Dr. Noriko Iwashita, who gave me insightful comments and advice. Her constant encouragement has helped me complete this thesis.

I would also like to thank my lecturers at UQ, Prof. Richard Baldouf Jr., Dr. Michael Harrington, and Dr. Kichun Na, who have shared their knowledge with me during my study.

I am grateful to all the research participants, including my teacher colleagues who helped arrange the field work for me.

Many thanks go to all my friends, Jessica, Orn, Sumaya, and Tina who gave invaluable advice and a lot of encouragement to me. I'd also like to include my friends in Korea to thank them for their spiritual support.

And finally, my sincere thanks go to my family. Without their great love and concern, I could not have completed this thesis.

ABSTRACT

The present study investigates the current practice in classroom speaking assessment in secondary schools in South Korea. Teacher-based speaking assessments conducted in the classroom are not only strongly recommended in Korean educational policies but are also the only tool used to evaluate students' oral skills in the formal schooling system. However, there has been little systematic research investigating *how* teachers *actually* assess students' oral skills in the classroom. Therefore, this study aims to investigate the current status of classroom speaking assessment in Korean middle schools and its effectiveness in light of an alternative assessment tool and pedagogical values. The data was collected from questionnaires and interviews where teachers were the only targeted respondents. 51 Korean English teachers recently working in middle schools participated in the questionnaire and six of them were interviewed. The results have revealed that classroom speaking assessment currently conducted in Korean middle schools has broadly employed performance-based tasks and that somewhat informative feedback has been offered to students in the form of criterion descriptions plus marking scores. However there was still a strong tendency here towards traditional formal testing to measure and report learning outcomes, one which resulted in teachers having an overall pessimistic attitude towards the positive effects of such testing on teaching and learning. It is evident from this study that there is need for improvements in order to facilitate better learning outcomes in the classroom. The study provides a range of suggestions for an improvement of current practices, starting with a process to change the perceptions of teachers, students, parents and policy makers towards classroom assessment followed by practical actions such as teacher training, cooperation with an English native teacher, and downsizing the number of students per class.

TABLE OF CONTENTS

STATEMENT OF ORIGINALITY	i
ACKNOWLEDGEMENTS	ii
ABSTRACT	iii
TABLE OF CONTENTS	iv
CHAPTER 1 INTRODUCTION	1
1.1. Overview	1
1.2 Research Background	3
1.2.1 English education in South Korean schooling system	3
1.2.2 Characteristics of English assessment in South Korean schooling system	7
1.3 Key term definitions	8
1.4 Outline of chapters	11
CHAPTER 2 LITERATURE REVIEW	12
2.1 Theoretical background	12
2.2 The advantages of classroom assessment	15
2.2.1 Pedagogical advantages	15
2.2.2 Advantages for speaking assessment	17
2.3 Issues related to reliability in classroom assessment	19
2.4 Conclusion	22
CHAPTER 3 METHODOLOGY	24
3.1 Participants	24
3.2 Data collection instruments	25
3.3 Data collection procedures	28
3.4 Data analysis	29

CHAPTER 4 RESULTS	34
4.1 Current status of classroom speaking assessment (RQ1)	34
4.1.1 Teacher profiles	35
4.1.2 RQ1a: What is the purpose of classroom speaking assessment?	35
4.1.3 RQ1b: What are the assessment practices in terms of assessment methods, frequency, and teacher feedback?	38
4.1.3.1 Assessment tasks	38
4.1.3.2 Frequency of assessment	42
4.1.3.3 Types of teacher feedback	43
4.2 Teachers' perceptions of the effectiveness of classroom speaking assessment (RQ2)...44	
4.2.1 Role of classroom speaking assessment	44
4.2.2 Positive effects of classroom speaking assessment on teaching and learning.45	
4.3 Conclusion	48
CHAPTER 5 DISCUSSION	49
5.1 Summary of the findings	49
5.2 Advantages and shortcomings of classroom speaking assessment	50
5.3 Difficulties in implementing/ conducting classroom speaking assessment	52
5.4 Conclusion	53
CHAPTER 6 CONCLUSION	54
6.1 Introduction	54
6.2 Implications	55
6.3 Limitations	56
6.4 Suggestions for further research	56
REFERENCES	58

APPENDIX A: QUESTIONNAIRE	65
APPENDIX B: INTERVIEW QUESTIONS	69
APPENDIX C: PARTICIPANT CONSENT FORM AND PARTICIPANT INFORMATION SHEET	70
APPENDIX D: TEACHER PROFILES	73
APPENDIX E: DESCRIPTIONS OF SPEAKING TASK FORMATS	74
APPENDIX F: SUMMARY OF INTERVIEWEES' BACKGROUND INFORMATION...	75
APPENDIX G: INTERVIEW TRANSCRIPTION SAMPLE	76

CHAPTER 1

INTRODUCTION

1.1 Overview

The present study investigates the current practice of classroom speaking assessment in secondary schools (i.e., Year 7 to 12) in South Korea. In Korea, an English language program is officially offered from Year 3 to 12, and three to four English classes are provided a week in secondary schools. Since the 6th National Curriculum was implemented in 1992, the curriculum has put an increasing emphasis on enhancing students' oral communication skills. However, there are few opportunities to evaluate students' speaking skills because speaking assessment is not administered in any formal exams including the high school entrance exam and the Korean version of the Scholastic Aptitude Test (KSAT)¹. Hence, unlike with the curriculum focus on this, in practice, English education in secondary schools puts a greater emphasis on passing the entrance exam.

Because of the Korean educational government's strong recommendation of teacher-based classroom assessment (via the Korea Institute of Curriculum and Evaluation) as well as curriculum emphasis on enhancing oral communication skills,

¹ However, Ministry of Education and Science in South Korea announced a plan of National English Ability Test (NEAT) administration in 2012 targeting students in the elementary school level to adults. The primary purpose of this test is to balance four language skills in the English curriculum by implementing direct testing of writing and speaking (KICE, 2009).

classroom assessment is widely practiced in many Korean secondary schools. For many students in Korea, classroom assessment is the only opportunity to have their speaking skills assessed unless they take external tests such as the Test of Spoken English (TSE), the Test of English as a Foreign Language (TOEFL), and the International English Language Testing System (IELTS).

Despite the government's strong recommendation of assessment by classroom teachers, there is still a challenge which stems from the emphasis for a measurement-driven orientation toward assessment. In other words, even if the rationale for classroom assessment is to provide both students and teachers with information about their current level of speaking skills and how to improve their speaking skills on the basis of their real language performance, teachers are nevertheless under pressure from such a policy to reflect the results of classroom assessment as a certain portion or percentage of the students' final grades. Based on the author's anecdotal evidence, it is not clear whether the initial intention of classroom assessment was appropriately implemented to satisfy the initial purpose of the assessment.

For that reason, the present study is to investigate what features classroom speaking assessment present in the situation where it is expected to achieve both pedagogic usefulness and meet the desired measurement purpose. Also, this study explores whether the current practice of classroom assessment in Korean secondary schools can be effective in terms of pedagogical benefits and speaking assessments. As Davison and Leung (2009) point out, there has been a lack of theorization of classroom

assessment in the English language teaching field, with researchers pointing to much variability, a lack of systematic principles and procedures, and a reliance on traditional psychometric testing. Moreover, there are few studies carried out in a Korean schooling context and if any, they are ones focusing on the elementary education level (i.e., Lee, 2007; Butler, 2009). Thus, this study seeks to provide information for secondary school teachers on the way to improve the delivery of their speaking assessments including the speaking instructions in the classroom.

The design of the current research study involves a combined quantitative and qualitative method. The questionnaire provides the quantitative data set concerning the details of the current practice of classroom assessment, and the qualitative method using the interview collects close-up information on teachers' perceptions of its effectiveness.

The following section presents the research background including the context of English education and characteristics of English assessment in the South Korean schooling system. Then, the definitions of the key terms are followed by the outline of chapters to follow.

1.2 Research background

1.2.1 English education in the South Korean schooling system

The history of English education in the South Korean schooling system began in 1946, consisting of eight stages, from the preparation stage to the 7th National Curriculum. The responsibility for designing the curriculum and developing all national

level examinations rests with the Korea Institute of Curriculum and Evaluation (KICE here after). During the 6th National Curriculum (1992-1999), English education was introduced at the elementary education level for the first time, and the focus of the curriculum was shifted from grammatical competence to communicative competence. At present (the 7th National Curriculum period ²), official English education is offered for ten years, four years at the elementary level and six years at the secondary school level. Also, the global goal of the English curriculum in the 7th National Curriculum is to enhance communicative competence, especially oral communication skills (KICE, 2005).

Another characteristic of the 7th National Curriculum is that particular curriculum objectives are proposed for each of elementary and secondary education. The general objective of English education in the elementary school level is to offer different types of curriculum—that is, intensive or complementary curriculum—on the basis of students' learning achievements. At secondary school, on the other hand, it aims at a level-based curriculum in which an extra supplementary course is provided for the lower achieving students.

In line with these two distinctive curriculum objectives, elementary and secondary educations also pursue different assessment objectives. The 7th National Curriculum clarifies the primary purpose of English assessment at the elementary

² The 7th National Curriculum has been applied to the Korean formal schooling system since 2000 (KICE, 2005).

education level: to encourage student's interest of English and motivate English acquisition, not to evaluate student achievement. Concerning such English assessment at the elementary education level, five objectives are outlined in the 7th National Curriculum as followings:

Assessment guidelines for the intensive and complementary curriculum

- 1) In the elementary school, assessment aims at a stimulation process which creates an atmosphere for students to pay more attention to learning.
- 2) Assessment is not associated with numeric scoring in order to reduce student inhibition.
- 3) It is recommended that the assessment offers written feedback by observing participation, attitudes, and communicative competence while involving communicative activities in the classroom.
- 4) Instruction and teaching mainly focus on speaking, and also assessment is carried out by observing oral skills and task involvement.
- 5) In classroom activities, grouping students based on their learning achievement should be deliberate.

(KICE, 2005, pp.158-159)

Hence, the assessment method is designed to offer individualized and informative feedback rather than merely numerical scores.

At the middle school level (equivalent to Year 7 to 9 in Australia), by contrast, the purpose of English assessment starts to be differentiated from the elementary education level as followings:

Assessment guidelines for the level-based curriculum

- 1) It is recommended that four language skills be assessed in balance, based on what has been learned in the classroom.
- 2) It is recommended that validity, reliability, and objectivity in assessment are improved.
- 3) It is recommended that discrete-point tests and integrative tests are used where appropriate.
- 4) To properly assess communicative competence, it is recommended that integrative testing be used
- 5) It is recommended to frequently check up on learning achievements and then complement student needs in order to produce fewer repeaters at each level.
- 6) It is recommended that assessment results are reflected in any teaching plan.

(KICE, 2005, p.159)

These guidelines, especially items 1 and 2 regarding balanced assessment of four language skills and valid assessments, may offer the foundation to the educational government's recommendation for implementing classroom assessment. This is because classroom assessment is prone to assessing writing and speaking skills based on student

language performance, one which is little assessed in the school-administered formal exam.

1.2.2 Characteristics of English assessment in South Korean schooling system

High stakes test situation

In Korean secondary education, English assessment plays an important role in determining whether students move to the next stage of schooling, that is, middle school to high school and high school to tertiary school. This accounts for a high stakes test situation in South Korea. High stakes testing refers to tests upon which test-takers' futures or study plans hinge due to its gate-keeping function (Davies et al, 1999). In other words, test outcomes are likely to affect the students' future careers or life directions. Because of the substantial impact of such assessment on students' futures and careers, students and parents are greatly concerned with their grades of English at school and the accurate judgment of this assessment work by their teachers. Therefore, under the high pressure from students and parents, Korean English teachers tend to develop less subjective testing in order to improve reliability and objectivity.

Assessing four discrete skills

The high stakes test situation influences the traditional exam-dominated culture in Korean secondary schools. For instance, in terms of the assessment format, pencil-and-paper tests prevail and the style is based on multiple choice questions where the correct response is sought to each. This test format may be seen as improving assessment

reliability and objectivity because as Hughes (2003) points out, no judgment is required on the part of the scorer so that scoring may be objective. At the same time, teachers are supposed to assess all four language skills in a balanced manner so as to satisfy the assessment guidelines as noted in the previous section. Consequently, teachers tend to assess the four skills separately. More specifically, conventional pencil and paper tests assessing reading comprehension including grammatical knowledge are the primary type of assessment, whereas listening, speaking, and writing assessments using performance-based tasks form a minor part of such assessment. Listening assessment is undertaken using the national level secondary school listening test administered once a semester by each province's educational council while speaking and writing assessments which do not have any authorized assessment tools, may be conducted as part of classroom assessment.

1.3 Key term definitions

Test vs. Assessment

In this study, the term *assessment* is distinguished from the term *test*. According to Brown (2004), *tests* refers to prepared administrative procedures that occurs at particular times in a curriculum in which learners manifest their competence for the best achievement, knowing that their responses are being measured and evaluated. Whereas used in a broader sense, *assessment* is an ongoing process of judgment, encompassing a teacher's comment and written phrase responding to students' performance as well as a

form of reporting measurement (Brown, 2004). That is, unlike tests, the results of assessment do not need to be reported all the time, and furthermore, they are not necessarily reflected in the final grades. Thus, the term *assessment* is used to incorporate a wide range of methods for evaluating student performance and attainment, including formal testing. Specific terms (i.e., standardized tests, formal tests, or school-administered tests) are used in cases where the discussion focuses on specific forms of assessment.

Classroom speaking assessment

Since there is little research specifically into *classroom speaking assessment*, this study makes reference to definitions of a few terms used interchangeably: *classroom-based assessment*, *teacher assessment*, and *alternative assessment*. According to Brown (2001), classroom-based assessment encompasses both teacher-designed formal and informal assessment in the context of day-by-day interaction with students, contrary to large-scale testing initiated by the school. It can be equivalent to teacher assessment (Underhill, 1987) which refers to tests designed by individual teachers (or a group of teachers) and administered in the process of classroom learning. According to Huerta-Marcías' (1995) claim, alternative assessment refers to alternatives to traditional testing. Thus, the characteristics of alternative assessment are summarized: (1) integrated in the curriculum; (2) provides information on the strength and weaknesses of each individual student; (3) provides multiple indices that can be used to gauge student progress; and (4) is more multiculturally sensitive and free of norms. Although such terms highlight

different aspects of the assessment process, all tend to share common concerns: a more teacher-mediated, context-based, classroom-embedded assessment practice, explicitly or implicitly defined in opposition to traditional externally set and assessed large scale formal examinations used primarily for selection and/or accountability purposes. Thus, for the purposes of this study, classroom speaking assessment is defined in such a way as to include *who* is doing the assessment, *what*, and *when*:

- The design of assessment methods and decision-making regarding evaluation should be driven by the students' own teacher;
- Regardless of the type of assessment method and its purpose, judgment should be made on students' speaking performance, that is, what they are doing, in contrast to some tests which seek to determine a student's knowledge about language;
- Assessment should be conducted during regular class time.

At this point, the extent to which the term *classroom speaking assessment* embraces the concept of performance tests needs to be determined. According to McNamara (2000) and Davies et al (1999)'s definitions, performance tests requires test takers to be involved in real language situations through communication or interaction, whereas in this study performance simply refers to all speaking activities, that is, what they do using oral skills.

1.4 Outline of chapters

Chapter 1 introduces the research background and provides definitions of key terms used in the study. Chapter 2 presents a review of the literature and relevant research with respect to the research problem and outlines the specific research questions addressed in the study. Chapter 3 describes the methodology used for data collection and analysis. Chapter 4 reports on the results of the data analysis. Chapter 5 provides a summary and a discussion of the research findings. Chapter 6 concludes the research with the implications for practice, limitation of the study and suggestions for further research.

CHAPTER 2

LITERATURE REVIEW

As discussed in the previous chapter, implementation of classroom assessment in Korean secondary schools is strongly recommended in Korean educational policies. This chapter presents the most relevant literature relating to the value of classroom assessment and draws on research concerning alternative assessment, classroom assessment, and communicative language testing. The first section provides a discourse on the constructivist learning theory and the socio-cultural theory to set the theoretical background. The next section highlights research on the advantages of classroom speaking assessment. Lastly, based on research findings, a discussion of issues relating to the reliability of classroom assessment is presented and conclusions drawn.

2.1 Theoretical background

In the language testing field, much of the research has been focused on external standardized tests which are used for making decisions about individuals and programs (Shohamy, 1994). However, as Rea-Dickins (2007) points out, with an increasing recognition of the significant limitations associated with a sole focus on learning outcomes as a measure of learner performance, the importance of the language learning process involving assessment has currently drawn more attention. The strength of classroom assessment lies in the fact that it is integrated with the normal processes of

teaching and learning and provides useful information on student learning in context (Leung, 2005). Thus, the distinctive usefulness and pedagogical value of classroom assessment can be taken into account in an alternative approach to the more conventional forms of standardized language testing. Its theoretical foundations, in turn, can be discussed in line with the concept of alternative assessment or constructivist learning theory and socio-cultural theory.

According to Gipps (1999), a constructive approach comes up with the postmodernist concept of denying the objectivity of knowledge which is unchangeable and everlasting. That is, in the perspective of constructivist learning theory, knowledge is actively sought through the social involvement and built up in an individual learner's cognitive process, and learning occurs by an active process of sense making (Shepard, 2000). In addition, Shepard (2000) explains how constructivists view both language and the learner's role in language learning. According to him, language is regarded as socially constructed and situated in a particular context of use rather than as an underlying trait or ability which remains stable across contexts. For language learning, learners as a human agent play a central role by actively making sense of new knowledge, making meaning from it, and mapping it into their existing knowledge map or schema. Thus, constructivists claim that contrary to standardized multiple-choice or short answer test which is efficient at sampling the acquisition of specific knowledge gained from a teacher, assessment should be able to examine in more depth the structure and quality of individual student's learning and understanding (Gipps, 1999).

These notions of constructivist learning theory draw on Vygotsky's socio-cultural theory. The key components of the latter theory are Vygotsky's (1978) *Zone of Proximal Development* (ZPD) which refers to the domain where the learner is not yet capable of independent functioning and *scaffolding* (Woods et al, 1976) or mediated support for cognition and communication. Accordingly, learners can reach the ZPD when they are given scaffolding, that is, the help of experts including teachers and more knowledgeable learners (Saville-Troike, 2006). Hence, learners' interaction with others (i.e. a teacher and peers) in their cultural, historical and institutional settings is an essential requirement for learning.

The constructivist learning theory and Vygotskian socio-cultural paradigm have implications for the theorization of assessment for learning. First, their claim of learner individual diversity provides the theoretical foundation to subjectivity and variability of assessment. In other words, assessment should offer equal opportunities for diverse learners to actively achieve learning outcomes depending on their unique and individual backgrounds rather than as they are consistently evaluated by predetermined assessment tools (Fox, 2008). Second, assessment should reflect a social context, considering that learners construct their knowledge and understandings within it (Shepard, 2000). Lastly, based on the concept of teacher-mediated assistance and the ZPD, assessment should allow the use of aids including teacher-mediated assistance or peer cooperation, to enable students to produce best performance (Gipps, 1999).

In summary, according to the perspectives of constructivist learning theory and

the socio-cultural theory, classroom assessment:

- is diverse
- employs authentic tasks reflecting an interactive language use such as used in the social and cultural life of the individual
- addresses learning process as well as learning outcomes
- is an ongoing process, integrated with instruction
- is used formatively in support of student learning

(Gipps, 1999; Shepard, 2000)

2.2 The advantages of classroom assessment

This section presents research findings on the advantages of classroom speaking assessment over traditional formal testing or external standardized tests in two aspects: pedagogical advantages and advantages for speaking assessment.

2.2.1 Pedagogical advantages

One major advantage of classroom assessment is that it can exploit its inherent feature, formative evaluation (Leung, 2005; SBA Consultancy Team, 2005; Davison & Leung, 2009). Bachman and Palmer (1996) suggest that the high feasibility of formative evaluation is largely due to the relatively fewer resources required for administering classroom assessment: for instance, a short vocabulary quiz and a sole teacher playing a role of an assessor. According to Brown (2004), the purpose of formative assessment is to evaluate students in the process of forming their competences and skills with the goal

of helping them to continue that growth process. Also, Davison and Leung (2009) claim that there are two key functions of formative evaluation—forming and informing—which bring benefits to the teachers as well as the learners.

For the teachers, formative evaluation allows them to form more reliable judgments in an ongoing process rather than making a quick decision based on a snapshot evaluation because here they are provided with sufficient opportunities to gather a great deal of information about knowledge, abilities and skills of the learners during classes (Underhill, 1987; Genesee & Upshur, 1996; Brown, 2001; Harris, 2007; Fulcher & Davidson, 2007). In addition, the information from formative evaluation is useful for teachers in planning their subsequent lessons in the teaching process, examining the effectiveness of their pedagogical objectives, and diagnosing student's strengths and weaknesses (Black & Wiliam, 2004).

Concerning the advantages to learners, feedback to such learners provided by formative assessment allows them to understand what they have learned and what they need to learn more (Black & Wiliam, 1998; Black, 2001; Black, Harrison, Lee, Marshall, & Wiliam, 2003; Brown 2004). In this regard, research to date has provided evidence of the importance of feedback formats. Stobart (2006) explains the vice of 'marks and grades only' feedback by citing Thorndike (1913) and other studies (Kohn, 1993; Reay and Wiliam, 1999; ARG, 2002). That is, grades can impede learning because they do not offer specific information but simply a comparison to others. In line with this, Butler (1988) investigates the effect of three different forms of

feedback—‘comments only’, ‘grades only’, and ‘a combination of grades and comments’—between two different ability groups. The study identified that lower achieving students expressed most interest towards the type of ‘comments only’ feedback. Byon (2005) and Gunn (1995) advance the argument of the effectiveness of criterion-referenced assessment in which the learner’s performance is interpreted by reference to predetermined criteria. According to them, unlike the summative tests primarily aimed at grading, marking criteria are informative and constructive regarding the extent to which learners can use language to achieve their communicative goals.

2.2.2 Advantages for speaking assessment

In fact, with respect to classroom assessment which targets oral skills, little research has been conducted. Most studies examine classroom assessment in general; for example, Dochy & McDowell (1997); Adamson & Davison (2003); Cumming & Maxwell (2004); and SBA consultant team, (2005). Thus, this section examines the advantages of classroom assessment for assessing oral skills by reference to research concerning performance assessment and communicative language tests as well as classroom assessment.

First, classroom assessment is more likely to offer opportunities for teachers to directly observe linguistic performance through classroom observations or direct testing (Brown, 2001). According to the Universal Grammar approach, linguistic competence—referring to the underlying linguistic abilities or knowledge of language—cannot be directly observed, whereas linguistic performance—the domain of language use—can be

directly observed (Mitchell & Myles, 2004). McNamara (1996) points out that the assessment of linguistic performance allows us to make direct inference of an individual's ability to use language appropriately or correctly in a variety of situations. Thus, a reliable and valid assessment method needs to evaluate a test-taker's performance rather than her/his linguistic knowledge. In particular, if the target language domain of assessment is *speaking*, then the importance of direct testing using performance tasks is critically increased. This is in line with Brown's (2001) argument that performance assessment contributes to improving content validity since students actually *perform* the target language use.

Also, Moon and Callahan (2001) insist that performance assessment can promote test authenticity because it presents learners with situations or asks them to perform academic exercises that simulate real-life experiences or problems. Test authenticity in relation to performance assessment can be considered in the communicative language teaching approach. In terms of Canale and Swain's (1980) communicative competence model, *speaking* by its nature needs to be judged on the basis of social context, interaction, communication, and integrated skills, all of which can be provided in the classroom context involving peers and a teacher.

Third, classroom assessment can serve as a powerful motivation factor. Crooks (1988) argues that tests in general have positive effects on enhancing learning motivation. According to Savignon (1997), especially in relation to communicative assessment, discrepancy between the test taker's grammatical competence and

communicative competence is manifested. Students can notice the gap between their linguistic competence and real language use while involving interaction with peers or a teacher. Consequently, it helps learners recognize what is really important and to know how to use what language they have learned in a variety of ways in real life situations.

Lastly, classroom assessment may be beneficial for reducing any student anxiety associated with test taking. Concerning the relationship between test anxiety and test results, both Phillips (1992) and Crooks (1988) provide evidence of a negative correlation between such anxiety and test achievements in the overall assessment situation. Crooks (1998) concludes that the negative influence of a learner's higher anxiety on achievement tends to be greater on standardized tests than in classroom assessment. Underhill (1987) explains the positive function of classroom assessment in reducing test anxiety in relation to the test-taker's familiarity with the test environment and the assessor. Namely, when assessed by the students' own teacher in the familiar classroom, students are more relaxed and confident and hence usually able to demonstrate better performance and proficiency than they can demonstrate in standardized tests.

2.3 Issues related to reliability in classroom assessment

Although classroom assessment has a number of advantages as presented in the previous section, the reliability issue remains controversial. In general, reliability of assessment is defined as consistency of measurement (Bachman & Palmer, 1996; David

et al, 1999; Brown, 2004). Thus, a reliable test is expected to be consistent across different characteristics of the testing situation. In the perspective of standardized tests, reliability is critical because the main purpose is to quantify the test-taker's ability for the purpose of public comparison. In this regard, classroom assessment may be less reliable than standardized tests due to the use of diverse tasks and the high involvement of teacher's judgment in evaluation. Birenbaum (1996), by citing empirical research found in Dunbar et al. (1991) and Linn et al (1991), claims that the use of diverse tasks in alternative assessment reduces the consistency in individual student performances across the different tasks.

Also, considering that rater's (i.e., teacher's) own judgment is highly involved in scoring, the reliability of classroom speaking assessment can be problematic such as occurs in other oral tests. Shohamy (1983)'s study provides evidence that the same test taker can be awarded different scores by different interviewers in oral proficiency interviews. As for the contribution of rater-oriented factors to undermining reliability, Underhill (1987) raises the possibility that the teacher's judgment may be influenced by relationships with learners over a period of time. Namely, teachers react in different ways to different students, and this may be reflected, if only subconsciously, in their assessment. Also, Llosa (2007) claims that teachers tend to have different abilities as an assessor, depending on background, training, and expectation. In addition, Brindley's (1998) study shows that rater severity makes major differences in test results concerning oral interaction, reading and writing.

Proponents of classroom assessment, on the other hand, take issue with the equal application of traditional reliability measures to classroom assessment. They argue that reliability of classroom assessment should be approached from a different perspective to the traditional standardized test, due to a conceptual difference which originates from the purpose of assessment: that is, assessment for learning vs. assessment of learning (Harlen, 2006); the interactive nature of student performance in the classroom vs. sampling one attribute as a property of the individual (Brown, 1995; McNamara, 1997); and a qualitative method vs. a quantitative method for objectively measuring at any one time (Lynch, 2001). Leung (2005) defines these qualities of classroom assessment as the low-stakes and home-made nature which helps bring about a better understanding of why we should offer concessions to classroom assessment from the psychometric perspective of high-stakes testing. In this regard, Gipps (1994) argues that since the level of *comparability* (i.e., reliability) demanded for any assessment will be related to its use, comparability in classroom assessment for formative purposes is of lesser concern. Haertel (1992) also suggests that a more relaxed attitude about reliability in classroom assessment helps to see the ways to improve reliability. According to him, through day-to-day classroom work, teachers generally have abundant opportunities and multiple sources of information about their students and furthermore, misinterpretation of teacher assessment outcomes can be easily corrected. Huerta-Marcías (1995) advances the stronger claim that the *trustworthiness* of alternative assessment is sufficient for claims of reliability. She accounts for reliability of alternative assessment

in aspects of *trustworthiness*: if the instrument and procedures of an assessment have *credibility* (i.e., truth value) and *auditability* (i.e., consistency), the assessment is *trustworthy*. Thus, alternative assessment which readily employs actual performances on real-life tasks provides valid data, and therefore it has a considerably high probability of producing consistently similar outcomes of a student's performance.

However, against Huerta-Marcías' (1995) argument, Kane, Crooks and Cohen (1999) argue that the relevance of the performance should be achieved without sacrificing too much reliability in spite of the context-dependent nature of classroom assessment, pointing out that if reliability is lost, the relevance of the performance is questionable because it cannot be easily measured.

2.4 Conclusion

This chapter has reviewed relevant research in the literature for the theoretical background, the advantages and the reliability in relation to classroom speaking assessment. Such research has conceptualized classroom assessment in the perspective of alternative assessment or the socio-cultural theory and identified its pedagogical usefulness and advantages to both teachers and learners in terms of formative evaluation. Also, different arguments about reliability in classroom assessment have been raised in the perspective of traditional testing vs. alternative assessment. With respect to advantages of speaking assessment in the classroom, on the other hand, they should be made with reference to performance assessment and communicative language tests

because there have been few studies focusing on speaking as a target language skill of classroom assessment. Moreover, there are even fewer studies based on the Korean secondary school experience. If any, a few studies (i.e., Lee, 2007; Butler, 2009) were conducted at the elementary education level. To date, research into classroom assessment has been carried out at a limited school level and has hardly focused on speaking. Therefore, this present study is intent on expanding the scope of this specific research field by exploring the secondary school level and investigating the detailed features of classroom speaking assessment. The following research questions will be addressed:

RQ1. What is the current status of classroom speaking assessment conducted in Korean middle schools?

RQ1a. What are the primary purposes of assessment?

RQ1b. What are the assessment practices in terms of assessment tasks, frequency, and teacher feedback?

RQ2. What are teachers' perceptions of the effectiveness of classroom speaking assessment?

CHAPTER 3

METHODOLOGY

This chapter describes the methodology of the current study involving teachers' responses to the current practice of classroom speaking assessment. Hence, to investigate the research problem, this study addressed two research questions: (1) *what is the current status of assessment regarding classroom speaking in South Korean middle schools?* and (2) *what are teachers' perceptions of its positive effects?* These two investigations involved a questionnaire and an interview respectively. First, a quantitative dataset was collected from the questionnaire portraying the overall picture of classroom assessment of speaking in the South Korean middle schools. Second, the data from teacher interviews was collated to provide in-depth information about individual teacher's perception of its benefits.

3.1 Participants

51 teachers aged over 22 who were currently working or had worked within the last two years in middle schools (i.e., Years 7 to 9 in Australia) in South Korea voluntarily participated in the research. Among those who indicated their willingness to participate in the interviews, six teachers actually participated in these follow-up interviews.

3.2 Data collection instruments

Questionnaires

The present study employed a questionnaire to investigate the first main research question (RQ1): *what the current status of assessment regarding classroom speaking conducted in South Korean middle schools is* and the two subsidiary questions: (RQ1a) *what the primary purposes of such assessment are* and (RQ1b) *what the assessment practices in terms of assessment tasks, frequency, and teacher feedback are*. According to Thomas (2003), the main strength of questionnaires is that they enable a researcher to provide a large quantity of factual information in a relatively short period of time. The format and content of the questionnaire partially replicated a previous study (Grierson, 1995) which was also interested in the nature and quality of classroom speaking assessment practice in the secondary education level in Australian Intensive English Centers. Yet, due to the different research interest and scope, a number of modifications were made as summarized below:

- It was written in both Korean and English to ensure participants' understanding of the questions.
- The questions enquiring about the purposes of assessment, methods of assessment, and assessment tasks were selectively drawn from Grierson (1995).
- Minor adjustments included implementing a Likert-scale according to the item purpose and changing some options pertinent to the research aim and context.

The questionnaire consisted of two sections. The first section had questions about

teacher profiles including gender, age, and length of teaching experience. In addition, the size of the class they taught was included.

The second section contained the two subsidiary questions mentioned above and sought to build on the information gathered in the first section above. The subsidiary questions were: (RQ1a) *what the primary purposes of classroom speaking assessments are* and (RQ1b) *what the assessment practices in terms of assessment tasks, frequency, and teacher feedback are*. There were three types of questions used here: closed, open-ended, and Likert-scale questions.

In RQ1a, two questions about current assessment purposes and the importance of assessment purposes as perceived by the teachers were asked: (Q3) *indicate in effect for what purpose you employ classroom speaking assessment* and (Q4) *indicate the importance of the purpose of classroom speaking assessment*. Teachers were encouraged to indicate one or more among the seven options in response to Q3 and to evaluate the importance of the assessment purposes as they perceived on a scale 1 to 5 in response to Q4.

In RQ1b about assessment practices, the three survey questions included assessment methods, specific speaking task formats, frequency, and format of teacher feedback as follows:

(Q5) *Which are the primary methods of classroom speaking assessment?*

(Q6) *Please write brief descriptions of up to 4 test tasks/ activities which you employed during the past year for your classroom assessment involving speaking*

and tick the given student grouping.

(Q7) Please choose one primary feedback that you would use below.

Because Q5 and Q7 were closed questions, teachers were asked to choose one option. In Q6, those who choose the option ‘marking specific speaking tasks’ in Q5 were then asked to give one to four short descriptions of task formats and to indicate the format for grouping students.

The questionnaire took approximately fifteen minutes to complete. Teachers were asked to complete the survey with their own nicknames to identify themselves in case of withdrawal. The questionnaire was piloted by two qualified English teachers prior to the main study. The full questionnaire is provided in Appendix A.

Interviews

The interviews were conducted to obtain further information about teachers’ perceptions regarding classroom speaking assessment in South Korean middle schools. A semi-structured interview format was selected on the basis of being the most appropriate for the research purpose, time availability, and requirements of flexibility. It allowed the researcher to ask a set of questions in a similar manner but also to maintain the flexibility to probe into relevant information where necessary.

The interview consisted of seven questions in three parts: (1) background information on teaching experience, teaching approach and assessment practices; (2) teacher’s perception of effectiveness of classroom speaking assessment; and (3) difficulties in conducting such assessment. The full interview questions are provided in

Appendix B.

3.3 Data collection procedures

The questionnaire was administered through online and face-to-face meeting. For the online survey, the following steps were undertaken. First, a recruiting advertisement was posted online at the website of the English teacher community in Korea (<http://www.njoyschool.net/>), requesting the participation of middle school teachers. Then, the Participant Consent Form and the Participant Information Sheet (see Appendix C) were e-mailed to those who agreed to participate in the survey. Next, once the signed consent form was sent to the researcher, the relevant participants were directed to the survey website (<http://www.surveymonkey.com/s/62W8ZMK>). Among the 42 English teachers who participated in the online survey, 25 fully completed responses were included in the data analysis.

In addition to the 42 teachers recruited online, another 30 teachers were invited from the researcher's personal contacts. Of the 30 teachers, 26 teachers agreed to participate in the survey. They were able to choose either the online survey tool or a face-to-face survey. All 26 teachers completed the survey. Consequently, responses from 51 participants in total were used for data analysis.

Interview participants were recruited from the survey respondents who had indicated their willingness to be interviewed. Then, those who had given full responses or additional comments in the questionnaire were contacted using contact details on the

questionnaires. In the end, six respondents participated in a face-to-face interview. All interviews were conducted at the quiet restaurant and each interview lasted approximately 30 minutes.

While all six interviews basically followed the same format and questions prepared prior to the interview, the researcher attempted to maintain the flow of natural conversation if the interviewee did not request any additional explanation on a question. All the interviews were conducted in Korean and audio-recorded with participants' consent.

3.4 Data analysis

Responses to the questionnaire and the interview data were analyzed to answer the first and second research questions respectively.

Questionnaires

The format of the current study is exploratory in nature. Three types of data (i.e., responses to open-ended and closed questions and responses to the Likert-scale question) collected from the questionnaire were analyzed in light of the two subsidiary questions (RQ1a and RQ1b). The frequency and/or the percentage were basically employed to describe the distribution of responses of closed and open-ended questions. The ratings in the Likert-scale were analyzed using descriptive statistics (see Chapter 4. Results). The following describes the specific procedures of data analysis according to each subsidiary question.

To answer RQ1a, “*what is the primary purpose of classroom speaking assessment?*” the responses to Q3 and Q4 were analyzed: the frequencies of responses to Q3 were calculated, and in Q4 teachers were encouraged to indicate the importance of their perceived purposes of assessment by choosing one rating from 1 to 5. The means and standard deviations were calculated.

Since RQ1b, “*What are the assessment practices in terms of assessment methods, frequency, and teacher feedback?*” contained three components of the assessment practice, the data from the Q5, Q6, and Q7 were analyzed according to each component. First, to explore the nature (and quality) of assessment methods, three aspects were considered: (1) the methods of classroom speaking assessment, (2) specific speaking tasks, and (3) the format of grouping students. Concerning (1): *the methods of classroom speaking assessment*, responses to Q5 were classified into seven categories. Three categories indicated choosing a sole method such as marking specific speaking tasks, observation, and peer assessment. The other four categories were a combination of two or more methods: for instance, marking specific speaking tasks and observation; marking specific speaking tasks, observation, and peer assessment; marking specific speaking tasks, observation and self assessment; and all four methods. Concerning (2): *specific speaking tasks*, 74 responses were provided from the open-end question, Q6. They were firstly divided into 13 categories according to similar characters of performance. For example, the descriptions of ‘role playing a model dialogue’ and ‘demonstrating a dialogue with a partner’ were

classified in the same category, that is, 'role play'. The 13 categories were then re-sorted into five types of speaking according to the taxonomy set out in Brown (2004). Brown's taxonomy classifies speaking activities into five types: *imitative*, *intensive*, *responsive*, *interactive*, and *extensive* (or *monologue*) types according to the extent of interaction, task allowance of open responses, and assessment purposes. The *imitative* type of speaking is the ability to simply repeat a word, phrases or a sentence. In the *intensive* type of speaking, test-takers are expected to demonstrate their comprehension of semantic properties as well as prosodic elements: for example, directed response tasks; reading aloud; sentence and dialogue completion; and limited picture-cued tasks. *Responsive* assessment tasks require interaction and test comprehension, and are limited in very short conversations. *Interactive* speaking is longer and complex than responsive speaking. Finally, *extensive* oral production tasks (i.e., *monologue*) include speech, oral presentation, and story-telling. With respect to (3) *the formats of grouping students*, the distribution of four options—individual, pair, small group, or whole class—was described as a frequency and a percentage.

As for the second component, the *frequency* of classroom speaking assessment, responses to Q5 were analyzed. Namely, the frequencies of each assessment method that individual teachers indicated in Q5 were added up, and then the total frequency of the speaking assessment that a teacher conducted in the classroom in a semester was described in the frequency and the percentage.

For the third component, the *format of teacher feedback*, the data from Q7 was

analyzed. Like the other closed questions, the distribution of the four types of feedback was described as a frequency and a percentage.

Interviews

All interviews were transcribed and translated into English by the researcher. As Marshall and Rossman (2006) indicate, translating is complex because it involves more subtle issues of connotation and meaning. Thus, the researcher sent the English transcription back to each interviewee via email to ensure that their intention and meaning of the interview were accurately and eloquently translated. If disagreements on transcriptions were raised, modifications were made. The analysis of the interview data followed the guideline in Akiyama (2004) as below:

- 1) Familiarization with the transcript;
- 2) Relating general units of meaning to the research focus;
- 3) Common patterns and themes extracted;
- 4) Re-analysis of data

The interview's purpose was to provide an answer to the second research question, "*What are teachers' perceptions of the effectiveness of classroom speaking assessment?*" The interview data was transcribed and analyzed according to the interview questions under the three themes: (1) background information on the length of teaching experience, teaching approach, and assessment practices, (2) teacher's perception of its effectiveness, and (3) difficulties in assessment administration. First the background information on the six teachers' teaching experiences, teaching

approaches, assessment methods, and a focused domain of assessment were summarized in a table (see Appendix F). Second, teachers' perceptions were categorized according to their attitudes: positive, negative, and neutral. For instance, teachers expressing the desire to keep conducting classroom speaking assessment in spite of some challenges were categorized into the positive group, whereas those who did not want to implement it fell into the negative group. Teachers who were not fully convinced of its useful role or positive effects even if they acknowledged its effectiveness to a degree were classified in the neutral group. Lastly, teachers' comments on the difficulties in assessment administration were analyzed along with the four teachers' responses to the closed question (Q2) about the reasons that they do not conduct classroom assessment. The frequent difficulties identified are provided in the discussion chapter.

CHAPTER 4

RESULTS

This chapter reports on the results of the questionnaire and the interview survey in light of the two main research questions.

RQ1. What is the current status of classroom speaking assessment conducted in Korean middle schools?

RQ1a. What are the primary purposes of assessment?

RQ1b. What are the assessment practices in terms of assessment methods, frequency, and teacher feedback?

RQ2. What are teachers' perceptions of the effectiveness of classroom speaking assessment?

The quantitative data collected from the questionnaire was to determine the overall status of classroom speaking assessment in the Korean middle schools (RQ1), and the interview elicited in-depth information of teachers' perceptions (RQ2).

4.1 Current status of classroom speaking assessment in Korean middle schools (RQ1)

For the first main research question, the profiles of the teachers will be outlined before presenting the results of the questionnaire on *the purposes and practices of classroom speaking assessment*. This is followed by a data analysis of the findings of the interview, incorporating the second research question RQ2: *What are*

teachers' perceptions of the effectiveness of classroom speaking assessment?

4.1.1 Teacher profiles

The majority of participants (approximately 86%) were female, and their ages mostly ranged from 31 to 45. 84% of the teachers had more than 5 years of teaching experience. A table summarizing the 51 teacher profiles is provided in Appendix D.

To investigate teachers' view of the benefits of classroom speaking assessment, first of all, it was asked whether the teachers had conducted speaking assessment in their classroom. Out of the 51 participants, 47 teachers (approximately 92%) indicated that they had conducted speaking assessment in their classrooms. Consequently, these 47 teachers' responses were included for data analysis. The account for not conducting the speaking assessments provided by the other four teachers will be incorporated into the discussion on difficulties in implementing/conducting classroom speaking assessment in the Korean secondary school context in the discussion chapter.

4.1.2 RQ1a: What is the purpose of classroom speaking assessment?

To examine this question, responses to two queries (Q3 and Q4) were analyzed.

Table 1 summarizes the results of the analysis of responses to Q3 (i.e., *for what purpose do you conduct classroom speaking assessment?*). In this question, the teachers were asked to choose one or more primary purpose(s) among the seven options. The most frequently chosen option was 'evaluation of a topic/unit of work' followed by 'to follow the curriculum policy of a school or the education council', and 'ongoing

assessment’.

Table 1: Primary purposes of classroom speaking assessment

Purpose	Frequency
Evaluation of a topic/unit of work	24
To follow the curriculum policy of a school or the education council	20
Ongoing student assessment (e.g. marking, feedback)	19
Ongoing programming for lesson planning	11
Pre-topic planning	10
Assessment of students for class placement	9
To provide information to others (e.g. bureaucratic report, parents, school)	7
Others	2

Note. The teachers had more than one choice.

‘To provide information to others (e.g. bureaucratic report, parents, and school)’ was marked as the lowest ranking answer. The comment noted by two teachers that classroom speaking assessment was used only for the purpose of summative evaluation was categorized in ‘others’.

In Q4, teachers were asked to evaluate the importance of each statement for assessment purposes using a five-point Likert scale (i.e., most important = 5, somewhat important = 4, neutral = 3, less important, and least important = 1). The descriptive statistics are displayed in Table 2.

Table 2: Mean importance of purposes of classroom speaking assessment (N=47)

Purpose	Mean	SD
a. To indicate students’ achievements of course objectives	3.86	0.89
b. To give students feedback on progress	3.85	0.94

<i>c.</i> To diagnose strengths and weaknesses and set further learning objectives	3.73	1.12
<i>d.</i> To give teachers feedback on learner's progress	3.5	1.17
<i>e.</i> To indicate levels of general speaking proficiency	2.96	1.02
<i>f.</i> To place students in a different level of the classroom	2.68	1.29
<i>g.</i> To provide information to others (e.g. bureaucratic report, parents, other teachers)	2.52	0.81

‘To indicate students’ achievements of course objectives’ was rated highest with a mean of 3.86, whereas the mean values of the top ranked four items were not much different considering the relatively large standard deviations. Thus, all the four purposes—that is, ‘to indicate students’ achievements’, ‘to give learners feedback on progress’, ‘to diagnose students’ strengths and weaknesses’, and ‘to give teachers feedback on learner’s progress’—showed the similar values as a perceived important purpose of classroom speaking assessment from teachers.

The overall results here showed a high correspondence with the results of the primary purposes of classroom speaking assessment (see Table 1). For instance, among the highly rated four items in the importance of assessment purposes, the three items (i.e., *a.* achievement tests, *b.* ongoing feedback to students, and *d.* ongoing feedback to teachers in Table 2) were also highly ranked in the primary purposes of the assessment. Likewise, the lowest rated two items—the purpose of class placement and providing information to parents or school (i.e., *f* and *g* in Table 2)—showed the lowest frequencies in the previous question as well.

4.1.3 RQ1b: What are the assessment practices in terms of assessment tasks, frequency, and teacher feedback?

To examine this question, the responses to three questions (Q5, Q6, and Q7) were analyzed.

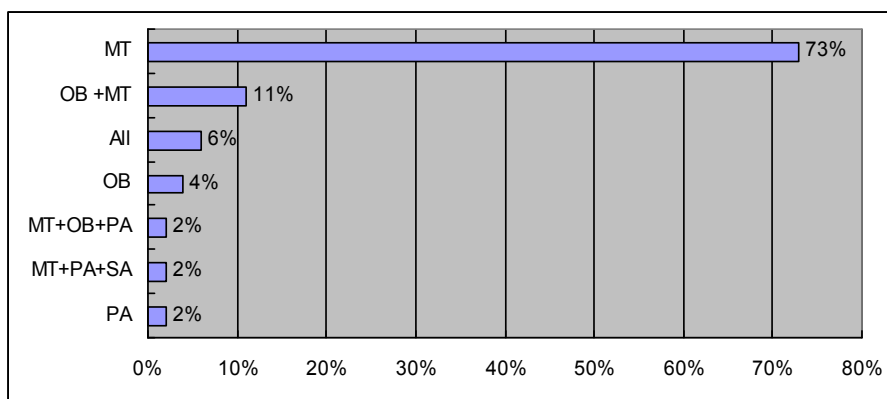
4.1.3.1 Assessment tasks

To gain results on the nature of assessment tasks, two sequential questions (Q5 and Q6) were analyzed in the aspects of assessment methods, specific task formats, and student grouping. In Q5, teachers were firstly asked to choose one or more methods for classroom speaking assessment. Then, in Q6, those who chose ‘marking specific speaking tasks’ in the previous question were asked to give brief descriptions of specific speaking task formats along with an indication of student grouping.

Methods of classroom speaking assessment (Q5)

In the teachers’ responses to the question, “*which method do you use for classroom speaking assessment?*” the majority chose ‘marking specific speaking tasks’ (see Figure 1).

Figure 1: Methods of classroom speaking assessment



Note. MT = Marking specific speaking tasks; OB = Observation; PA = Peer assessment; SA = Self assessment

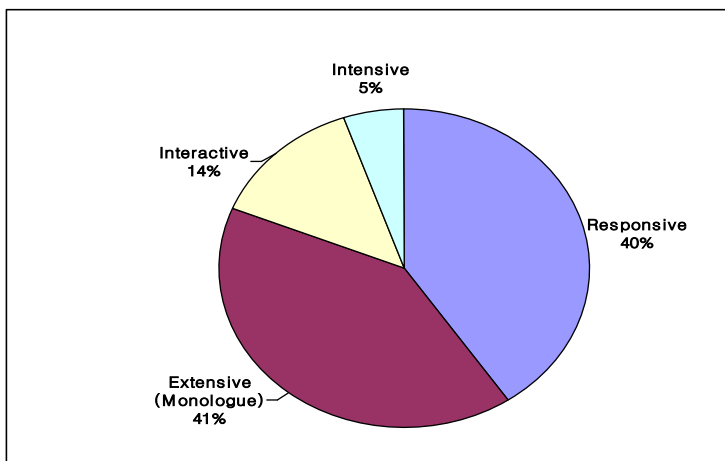
Moreover, adding up the cases where it was used in combination with other methods

(i.e., observation, peer assessment, and self-assessment), it accounted for approximately 94% of the methods. The ‘observation’ method, by contrast, was used along with other methods rather than solely used.

Speaking task formats (Q6)

Concerning the format of speaking tasks and grouping of students, a total of 74 responses were provided by 44 teachers who chose ‘marking specific speaking tasks’ in the previous question. They described at least one or up to four types of speaking tasks frequently used in their assessment. The results are summarized in Figure 2 and Table 5. Figure 2 displays the distribution of the four speaking types. As mentioned in the methodology chapter, categorization into these four types following Brown’s (2004) taxonomy, which describes the basic types of speaking with five categories such as *imitative*, *intensive*, *responsive*, *interactive*, and *extensive*. In the present study, however, the *imitative* type was not displayed because no task format fell into this category.

Figure 2. Categories of speaking task formats



The *responsive* and the *extensive* were predominant categories, indicating approximately 40% (see Figure 2). In contrast, the *interactive* and the *intensive* category presented relatively less popularity.

Table 3 shows the frequently used speaking task formats in each category.

Table 3. Specific speaking task description

Category	Frequently used speaking task format	Frequency	Task description
(1) Responsive	Role-play	12	In pairs demonstrate a model dialogue based on memorization
	Question and answer	8	Answer a set of questions given prior to the assessment
(2) Extensive (Monologue)	Speech	11	Demonstrate a written script based on memorization
	Reciting	10	Recite useful sentences
(3) Interactive	Interview	7	Interview between a teacher and a student with preplanned interview questions
(4) Intensive	Reading out	3	Reading aloud sentences, reading out student's own classroom reflection, and reading a passage in the textbook

Note. Here the frequently used speaking task formats of each category were presented among a total of 74 responses. The full speaking task formats can be reviewed in Appendix E.

In the *responsive* category, ‘role-play’ was the most frequently used task. Out of 18 responses indicating ‘role-play’ format, the majority (i.e., 12 responses) were based on memorizing a dialogue in the textbook. Also, in the ‘question and answer’ format, students could rehearse what they would answer in the assessment because a list of questions was suggested to students beforehand.

In the *extensive (monologue)* category, ‘speech’ and ‘reciting’ were the frequently used speaking tasks. Both task descriptions showed that students’ spontaneous responses were considerably controlled. As for ‘speech’ tasks, all the 11 teachers described the tasks as being adapted from a reciting format by allowing students to memorize a written script. In a similar vein, in the ‘reciting’ task format, students demonstrated or recited materials that they had memorized such as an English pop song or useful expressions provided prior to the assessment. Consequently, memorizing-based task formats including ‘role play’, ‘speech’, and ‘reciting’ were broadly used giving a total response of 52.64%. Moreover, in terms of the limited possibility of spontaneous responses, approximately 64% of responses including the ‘question and answer’ format as well as ‘role play’, ‘speech’, and ‘reciting’ were based on the preset situation.

Interestingly, out of all 74 cases, merely two task formats in the ‘intensive’ category—‘reading aloud sentences’ and ‘reading a passage in the textbook’—were the type of tasks that assess phonological facets including intonation, pronunciation, stress, rhythm, and junction. The result may suggest that evaluation of phonological features is rarely aimed at in classroom speaking assessment.

Types of student grouping (Q6)

For supplementary information to describe speaking task formats, teachers were asked to indicate one of the four types of student grouping required in performance on the assessment. Table 4 summarizes the results.

Table 4: Types of student grouping

Student grouping	%	Frequency
Individual	46	34
Pair	48	35
Small group	5	4
Whole class	1	1
Total	100	74

‘Pair’ was chosen as the most frequently used format of student grouping, while the percentage of ‘individual’ (46%) was almost as high as the one of ‘pair’ (48%). This result provided information to infer the interactive quality of assessment tasks.

To summarize the findings above, the majority of teachers adopted specific speaking tasks for marking, and the task formats presented these distinctive features: preplanned responses and less involvement of interpersonal exchanges.

4.1.3.2 Frequency of assessment

The frequencies indicated in the individual four assessment methods in Q5 were added up, and then the total frequency of classroom speaking assessment that a teacher conducted a semester were summarized in Table 5.

Table 5: Total frequency of classroom speaking assessment

Total frequency (a semester)	%	Frequency
Once	54	25
Twice	22	10
Three	6	3
Four	6	3
Five	4	2
Six	6	3
More than six	2	1
Total	100	47

54% of respondents indicated that classroom speaking assessment was carried out once a semester, and furthermore, ‘once’ and ‘twice’ prevailed with 76%.

4.1.3.3 Types of teacher feedback (Q7)

To examine the type of teacher feedback, four options were provided: *marking scores*, *criterion descriptions plus marking scores*, *oral feedback*, and *written feedback*.

Table 8 summarizes the results. 64% of teachers chose ‘criterion descriptions plus marking scores’, while 28% of teachers responded to provide solely ‘marking scores’.

Although the percentage of ‘marking scores’ (i.e., 28%) was relatively lower than the one of ‘criterion descriptions plus marking scores’, the difference of the percentage of the oral or the written feedback (respectively 6% and 2%) was distinctive.

Table 8. Types of teacher feedback

Feedback	%	Frequency
Criterion descriptions plus marking scores	64	30
Marking scores	28	13
Oral feedback	6	3
Written feedback	2	1
Total	100	47

4.2 Teachers' perceptions of effectiveness classroom speaking assessment (RQ2)

This section primarily reports on the results of the analyses of the interview data. The two questions asked relating to teachers' perceptions of effectiveness of classroom speaking assessment were: *what they thought the role of classroom speaking assessment was, and what positive effects of classroom speaking assessment they perceived on teaching and learning were*. Six teachers had approximately average seven year teaching experience in secondary schools, and the main teaching approach was a grammar-translation method. They used the *responsive*, the *extensive (monologue)*, and the *interactive* type of speaking tasks as similarly as the frequently used task formats revealed from the questionnaire survey (see Table 3). Further details are provided in Appendix F.

4.2.1 Role of classroom speaking assessment

Three of six teachers showed positive attitudes towards classroom speaking assessment as indicated by the comments involving *benefits for the students' confidence, learning motivation, class participation and so on*:

- (1) *Assessment is an effective tool to encourage learning motivation especially in a short term. Without assessment, it would be difficult to get students eagerly involving speaking activities.* (Teacher B)

Teacher O's belief that assessment can encourage students' participation during the regular class time was in accordance with her assessment practices. Namely, she commented that her only assessment criterion was a student's attempt to use English.

Thus, she gave marks if a student participated in class activities. Teacher H said that classroom assessment may contribute to reducing students' inhibition caused by their low proficiency. She maintained that students as a result could have more opportunities not only to practice oral skills while preparing the assessment but also to speak out in front of many people. She expected that eventually, her students could somewhat build up their confidence.

However, the other three teachers had a negative outlook regarding the role of classroom speaking assessment, in particular, that the current speaking assessment methods hardly allow teachers to assess students' genuine speaking competence. Teacher Y made the strongest claim of all as shown below:

(2) I think that the current classroom speaking assessment does not play any important role to facilitate speaking or to motivate learning. Currently, it just serves to mark students' instant memorizing abilities and to report the outcomes. Moreover, it creates a great deal of assessment stress to students.(Teacher Y)

4.2.2 Positive effects of classroom speaking assessment on teaching and learning

Only two teachers mentioned any positive effects of classroom speaking assessment on their teaching, the other four teachers not perceiving any evident effects. The first two teachers commented that classroom speaking assessment was beneficial for speaking instructions and organizing the next teaching plan. That is, Teacher B said that she employed a greater variety of communicative activities apart from the structured curriculum based on the textbook, in order to conduct speaking assessment.

The reason given was that assessment should be in line with what students have already learned during classes. Also, Teacher P stated that she used general feedback about students' weaknesses provided by the classroom speaking assessment to set up her next teaching plan.

In terms of learning, teachers' attitudes can be divided into three categories: positive, neutral, and negative. The majority of interviewees, four teachers, had a somewhat neutral position, and the rest of two teachers revealed respectively positive and negative opinions about its effectiveness for learning. Only Teacher B perceived positive effects on students' motivation towards learning, stating "classroom speaking assessment may increase students' recognition of how important speaking is in learning English. Although the (speaking) assessment is not frequently conducted, I think that even one or two opportunities in a semester could address the requirements of learning English."

All of the four teachers who had a neutral position commonly indicated the limitations of the current methods of classroom speaking assessment. Teacher H and Teacher S mentioned that with the current method based on memorization, it would be hard to see positive effects on learning in the short term, but it may be beneficial to some students' real language uses in the long term. In this regard, Teacher O who commented on its positive function in encouraging students' participation during the class was not fully convinced of its direct effect on learning:

(3) *I think that classroom speaking assessment is meaningful because it can offer opportunities for students to use English. Due to having less confidence and passiveness, my students are not willing to speak out. If they should say something in English for assessment, they may realize that they can do it. In terms of speaking achievement, however, I am not sure that such assessment has evident positive effects.*(Teacher O)

Teacher P was even less convinced, stating “with respect to the current assessment tasks, it may facilitate linguistic competence including grammar and vocabularies. However, those domains can be assessed by other types of assessment, not merely by speaking assessment. Thus, I do not want to implement speaking assessment in my classroom if possible.”

Teacher Y was most pessimistic. Her claim stemmed from the English learning environment in Korea, that is, an English as a Foreign Language (EFL) situation. She commented:

(4) *I am doubtful about the effectiveness of classroom speaking assessment. Most students easily forget what they learned during assessment without reinforcement. As mentioned earlier, they are provided with few opportunities to enhance what they learned either in the classroom or in the real life.* (Teacher Y)

To sum up, the teachers were somewhat consistent in their attitudes towards the positive effects of classroom speaking assessment with their attitudes towards its role. However, even the teachers who commented on the useful roles of classroom speaking assessment, showed skepticism about its positive effects on teaching or learning. Most teachers’ pessimistic attitudes stemmed from the current assessment methods which

were characterized by a lack of spontaneous responses and interpersonal exchanges.

4.3 Conclusion

This chapter has reported on the results of the research to answer the two main research questions: *the current status of classroom speaking assessment in Korean middle schools* and *teachers' perceptions of its effectiveness*. In the following chapter, the findings in this study will be summarized and discussed.

CHAPTER 5

DISCUSSION

5.1 Summary of the results

The previous chapter reported on the results of the questionnaire and the interview.

The findings are summarized in light of the research questions:

RQ1. What is the current status of classroom speaking assessment conducted in Korean middle schools?

RQ1a. What are the primary purposes of assessment?

RQ1b. What are the assessment practices with respect to assessment methods, frequency, and teacher feedback?

RQ2. What are teachers' perceptions of the effectiveness of classroom speaking assessment?

The results revealed that the primary purposes of classroom speaking assessment were to evaluate a topic/unit of work and to follow requirements of Korean educational policy. Also, the study found that classroom speaking assessment was mostly conducted in the form of testing using speaking performance tasks and held once or twice in a semester, and that most frequent teacher feedback to students was criterion descriptions plus marking scores. With respect to teachers' perceptions, overall, teachers showed a pessimistic attitude towards its positive effects on teaching and learning due to current assessment practices. On the basis of the results, there is some evidence to suggest that

the nature of classroom speaking assessment in Korean middle schools has both advantages and shortcomings in terms of a valid speaking assessment strategy and pedagogical usefulness. These issues will be further discussed in the following section.

5.2 Advantages and shortcomings of classroom speaking assessment

First, the current speaking assessment methods seem to attain test authenticity to some extent. As McNamara (1996) points out, direct testing based on students' performance is more likely to facilitate an inference about a student's oral ability to use language in real life situations than would the indirect assessment method. In this regard, performance based assessment can improve the authenticity of language assessment. The results in Section 4.1.3.1 revealed that the majority of teachers conducted classroom speaking assessment using performance-based speaking tasks.

Second, informative feedback is offered using criterion descriptions. As Byon (2005) and Gunn (1995) claim, marking criteria have informative and constructive functions by indicating the extent to which learners can use language to achieve their communication goals.

This study, however, found that classroom speaking assessment in Korean middle schools had a strong tendency towards traditional testing. The results in Section 4.1.2 and 4.1.3 showed that the primary purposes of assessment were for summative evaluation and accountability requirements; that the speaking assessment tasks presented controlled features such as preplanned responses and less involvement of

interpersonal exchanges; and that teacher feedback mainly took the form of marking scores.

The above mentioned controlled features of assessment tasks undermine content validity because there is the incongruity between target language use and speaking skills actually assessed. This potential pitfall is also identified in Rea-Dickins and Rixon's (1999) case study. Their study showed that unlike the pedagogical approach, speaking skills actually assessed were rehearsed dialogues with little or no opportunity for spontaneous language use.

The reason for this traditional testing tendency can be found in relation to the issue of test reliability and the high stakes test situation of Korean secondary schooling. Namely, classroom speaking assessment is assigned to support learning, but at the same time it should be a reliable and objective measurement tool. For this, there is some evidence from the teacher interviews. The majority of teachers indicated that their main concern with regard to classroom speaking assessment was improvement of assessment reliability and objectivity. One teacher commented:

(5) Even if I know that classroom speaking assessment can benefit my students in quite a few ways, it was also important for me to develop a reliable and objective test considering that the outcomes of classroom speaking assessment were reflected in the final academic results which had a considerable influence on the decision towards high school admission. (Teacher P)

Another limitation is the low frequency of assessment administration for ongoing

assessment. On the basis of the results in Section 4.1.3.2, the average speaking assessment frequency carried out was 2.04 per a semester (the option, ‘more than six’, was excluded due to an exceptional case). Namely, taking account of the five-month semester in the Korean secondary schools, it does not seem that approximately two assessments per semester can be regarded as an effective ongoing process. Thus, unlike its primary purpose as reported by teachers, in practice it appears to be difficult to accumulate information or to frequently and regularly provide feedback about students’ learning during the progress of classroom speaking assessment. Concerning the reasons for the discrepancy between perceptions and real practices, the particular difficulties in conducting/implementing speaking assessment in the classroom will be discussed in the next section.

5.3 Difficulties in implementing/conducting classroom speaking assessment

The results on difficulties in implementing/conducting classroom speaking assessment were contributed to the responses of four teachers who did not conduct classroom speaking assessment and the interview data. The most challenging factor was an oversized class problem. Teachers argued that more than 40 students were too large a number to complete assessment within a class period of 45 minutes. According to the data on class size in the questionnaire, in the majority of cases a class had more than 36 students, and furthermore, in over one third of the cases, there were more than 41 students in a class (see Appendix D). One teacher remarked:

(6) *I prefer to complete a test within one class period for test fairness. This is because if a class takes the same test over more than one period, some students in the class would have more time to prepare for it.* (Teacher S)

In a similar vein, the Grierson's (1995) study identified that the most frequent cause for teachers' concern was lack of time caused by large numbers in a class. Also, another problem raised was students' low proficiency and heavy workload needed to prepare for such assessment.

5.4 Conclusion

Overall, the study of the current status of classroom speaking assessment in Korean middle schools pointed to both the qualities of classroom assessment as alternative assessment and the qualities of traditional formal examinations for testing and measuring students' achievements. On the other hand, the strong tendency of the testing administration in South Korea towards preferring traditional formal tests was the main reason for teachers' pessimistic perceptions of its positive effects on teaching and learning.

CHAPTER 6

CONCLUSION

6.1 Introduction

This study investigated the current practice of classroom speaking assessment in South Korean secondary schools by examining its current status and teachers' perceptions on its effectiveness. The questionnaire and interview targeted English teachers who were working in middle schools. The questionnaire drew out data which could portray the current status of assessment including purposes and practices, the latter encompassing the methods, frequency, and teacher feedback. The results of the data analysis show that classroom assessment was broadly conducted using specific speaking tasks and that criterion description plus marking scores were the main types of teacher feedback. This suggests that classroom speaking assessment currently conducted in Korean middle schools had the possibility of being an authentic tool in terms of being both a benchmark speaking assessment style and a supportive learning strategy with informative feedback. Still, it presented a strong tendency towards traditional formal testing for measurement and reporting learning outcomes. Although this tendency seems to stem from the need for measurement and testing, it is evident from this study that this system of assessment needs improvement in order to facilitate more effective teaching and learning. In this regard, this study has implications for the way to improve such classroom speaking assessment.

6.2 Implications

The implications of this study can be suggested in the realms of both perceptions and practices. To solve the problem caused by the dominance of summative assessment requirements, above all, there may be the need for stakeholders to change their perceptions of assessment: from being a tool for measurement to one for learning. In fact, as Fullan (2001) mentions, changes in individual beliefs are even more difficult than implementation of the new policy or an innovation. Teachers frequently appeal, that they have no choice due to the situational constraints, and hence a strategy to change perceptions should first target not only the teachers but also other test users including students, parents, and policy makers. Eventually, this can lead to an alteration in teaching practice such that *learning* becomes the most important goal of assessment.

Next, we should take account of a way to improve test reliability in spite of the context-dependent nature of classroom assessment. In this regard, improved teacher training courses can work to foster the qualification and professionalism of teachers as assessors and also improve their English ability—the challenge inevitably embedded with non-native teachers. With respect to a way of promoting test reliability, another suggestion is for cooperation of non-native English speaking teachers with native-English speaking teachers in the evaluation aspect of students' assessment. The linguistic intuition of the 'native' speaker must be used to assist 'non-native' teacher's evaluation here.

Lastly, greater funding for improved school infrastructure should be provided in

order to facilitate effective classroom assessment as the primary purpose of such improvements: for example, more classrooms built to enable a reduction in the large class sizes, supplying more assessment aid tools for video/audio recording and more native English teachers to assist teaching and assessing oral skills.

6.3 Limitations

There are some limitations in this study. The first possible limitation resides in the choice of participants. The lack of regional balance may not be sufficiently representative of the whole teaching population.

The second limitation is related to the questionnaire. As Thomas (2003) indicates, unlike interviews, questionnaires rarely provide opportunities for participants to receive any clarifications of confusing items. Due to this, the wording of some of the questions might cause some misunderstandings for respondents.

Lastly, the absence of the students' view of the effectiveness of classroom speaking assessment remains a limitation of this study due to the difficulty in recruiting young people.

6.4 Suggestions for further research

For further research there is a need to extend the scope of the investigation and to select a range of teachers providing a regional balance. By including students' perceptions in the study along with those of teachers and also balancing teacher participants across the nation, a more complete picture of the purposes and practices of

classroom speaking assessment in Korean secondary schools can be drawn and also an improvement in the level of confidence given to the authenticity of the results, effected.

REFERENCES

- Adamson, B., & Davison, C. (2003). Innovation in English language teaching in Hong Kong primary schools: One step forwards, two steps sideways. *Prospect*, 18, 27-41.
- Akiyama, T. (2004), *Introducing EFL speaking tests into a Japanese senior high school entrance examination*. Doctoral dissertation, The University of Melbourne.
- ARG. (2002). *Testing, motivation, and learning*. University of Cambridge: Assessment Reform Group.
- Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice: Designing and developing useful language tests*. Oxford, [England]: Oxford University Press.
- Birenbaum, M. (1996). Assessment 2000: Towards a pluralistic approach to assessment. In M. Birenbaum, & F. Dochy (Eds.), *Alternatives in assessment of achievements, learning processes and prior knowledge (pp.3-29)*. Boston, Mass.: Kluwer Academic Publishers.
- Black, P. (2001). Formative assessment and curriculum consequences. In D. Scott (Ed.), *Curriculum and assessment*. Westport, CT: Ablex.
- Black, P., & Wiliam, D. (1998). Assessment and Classroom Learning. *Assessment in Education* 5(1), 7-74.
- Black, P. & Wiliam, D. (2004). The formative purpose: Assessment must first promote learning. In M. Wilson (Ed.), *Towards coherence between classroom assessment and accountability (pp.20-50)*. Chicago: National Society for the Study of Education.
- Black, P., Harrison, C., Lee, C., Marshall, B. and William, D. (2003) *Assessment for learning: Putting it into practice*. Buckingham: Open University Press.
- Brindley, G. (1998). Assessing in the AMEP: Current trends and future directions. *Prospect*, 13(3), 59-73.

- Brown, A. (1995). The effect of rater variables in the development of an occupation-specific language performance test. *Language Testing*, 12(1), 1-15.
- Brown, H. D. (2001). *Teaching by principles: An interactive approach to language pedagogy* (2nd ed.). New York: Addison Wesley Longman, Inc.
- Brown, H. D. (2004). *Language assessment: Principles and classroom practices*. New York: Pearson/Longman.
- Butler, R. (1988). Enhancing and undermining intrinsic motivation: the effect of task-involving and ego-involving evaluation on interest and performance. *British Journal of Educational Psychology*, 62, 934-943.
- Butler, Y. (2009). How do teachers observe and evaluate elementary school students' foreign language performance? A case study from South Korea. *TESOL quarterly* 43(3), 417-444.
- Byon, A. S. (2005). Classroom Assessment Tools and Students' Affective Stances: KFL Classroom Settings. *Language and Education*, 19(3), 173-193.
- Canale, M., & Swain, M. (1980). *Approaches to communicative competence*. Singapore: Seameo Regional Language Centre.
- Crooks, T. (1988). The impact of classroom evaluation practices on students. *Review of Educational Research*, 58 (4), 438-481.
- Cumming, J., & Maxwell, G. (2004). Assessment in Australian schools: Current practice and trends. *Assessment in Education*, 11, 89-108.
- Davies A., Brown A., Elder, C., Hill, K., Lumley T., & McNamara, T. (1999). *Dictionary of language testing*. Cambridge; New York: Cambridge University Press.
- Davison, C., & Leung, C. (2009). Current issues in English language teacher-based assessment. *TESOL quarterly*, 43(3), 393-415.

- Dochy, F., & McDowell, L. (1997). Assessment as a tool for learning. *Studies in Educational Evaluation*, 23(4), 279-298.
- Dunbar, S., Koretz, D., & Hoover, H. (1991). Quality control in the development and use of performance assessments. *Applied Measurement in Education*, 4, 290-303.
- Fox, J. (2008). Alternative assessment. In E. Shohamy and N. H. Hornberger (eds.), *Language testing and assessment* (2nd ed., pp.97-108). New York: Springer.
- Fullan, M. (2001). *The new meaning of educational change* (3rd ed.). New York: Teachers College, Columbia University.
- Fulcher, G., & Davidson, F. (2007). *Language testing and assessment: an advanced resource book*. Abingdon; New York: Routledge.
- Genesee, F., & Upshur, J. A. (1996). *Classroom-based evaluation in second language education*. Cambridge, UK; New York: Cambridge University Press.
- Gipps, C. (1994). *Beyond testing: Towards a theory of educational measurement*. London: Falmer Press.
- Gipps, C. (1999). Socio-cultural aspects of assessment. *Review of Research in Education*, 24, 355-392.
- Grierson, J. (1995). Classroom-based assessment in Intensive English Centers. In G. Brindley (Ed.), *Language assessment in action* (pp.195-238). NSW: National Centre for English Language.
- Gunn, M. (1995). Criterion-based assessment: a classroom teacher's perspective. In G. Brindley (Ed.), *Language assessment in action* (pp.239-270). NSW: National Centre for English Language.
- Haertel, E. (1992). Performance measurement. In M.C.Alkins (Ed.), *Encyclopedia of Educational Research* (6th ed., Vol.13, pp.984-989). New York: Macmillan.

- Harlen, W. (2006). The role of assessment in developing motivation for learning. In J. Gardner (Ed.), *Assessment and learning* (pp.61-80). Thousand Oaks, CA: Sage Publications.
- Harris, L. (2007). Employing formative assessment in the classroom. *Improving Schools*, 10(3), 249-260.
- Huerta-Marcías, A. (1995). Alternative assessment: Responses to commonly asked questions. *Tesol Journal*, 5(1), 8-11.
- Hughes, A. (2003). *Testing for language teachers* (2nd ed.). Cambridge; New York: Cambridge University Press.
- Kane, M., Crooks, T., & Cohen, A. (1999). Validating measures of performance. *Educational Measurement: Issues and Practice*, 18(2), 69-95.
- Kohn, A. (1993). *Punished by rewards: The trouble with gold stars, incentive plans, A's, praise, and other bribes*. Boston: Houghton Mifflin.
- Korea Institute of Curriculum and Evaluation. (2005). *Youngeogoa Kyouytkkoajeoung Kaesonbanganeongoo*. Korea: The Ministry of Education and Science. Retrieved March, 3, 2010 from http://www.kice.re.kr/ko/board/view.do?article_id=49468&menu_id=10077&search_field=title&search_key=%BF%B5%BE%EE
- Korea Institute of Curriculum and Evaluation. (2009). *Malhagi & sseugi neungryeok sinjangeul euihan kugayeoungoneungryokpyeonggasiheom geybal bangan*. Korea: The Ministry of Education, Science and Technology.
- Lee, W.K. (2007). Assessment used in teaching English as a foreign language at elementary schools in Asia: Korea's case. *Paper presented at the 2007 Asia TEFL International Conference*. Kuala Lumpur, Malaysia.
- Leung, C. (2005) Classroom teacher assessment. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp.869-888). Mahwah, N.J.: L. Erlbaum Associates.

- Linn, R., Baker, E., & Dunbar, S. (1991). Complex, performance-based assessment: Expectations and validation criteria. *Educational Research*, 16, 15-21.
- Llosa, L. (2007). Validating a standards-based classroom assessment of English proficiency: A multitrait-multimethod approach. *Language Testing*, 24(4), 489-515.
- Lynch, B. (2001). Rethinking assessment from a critical perspective. *Language Testing*, 18(4), 333-349.
- Marshall, C., & Rossman, G. B. (2006). *Designing qualitative research* (4th ed.). Thousands Oaks, Calif.: Sage Publications.
- McNamara, T. (1996). *Measuring second language performance*. London; New York: Longman.
- McNamara, T. (1997). 'Interaction' in second language performance assessment: Whose performance? *Applied Linguistics*, 18(4), 446-466.
- McNamara, T. (2000). *Language testing*. Oxford: Oxford University Press.
- Mitchell, R., & Myles, F. (2004). *Second language learning theories* (2nd ed.). London; New York: Arnold; Distributed in the United States of America by Oxford University Press.
- Moon, T. R. & Callahan, C. M. (2001). Classroom performance assessment: What should it look like in a standards-based classroom? *NASSP Bulletin*, 85(622), 48-58.
- Phillips, E. (1992). The effects of language anxiety on students' oral test performance and attitudes. *The Modern Language Journal*, 76(1), 14-26.
- Reay, D., & Wiliam. D. (1999). I will be a nothing: structure, agency and the construction of identity through assessment. *British Educational Research Journal*, 25, 343-354.
- Rea-Dickins, P. (2007). Classroom-based assessment: Possibility and pitfalls. In J. Cummins, & C. Davison (eds.), *International handbook of English language teaching* (pp.505-520). New York: Springer.

- Rea-Dickins, P., & Rixon, S. (1999). Assessment of young learners' English—Reasons and means. In S. Rixon (Ed.), *Young learners of English: Some research perspectives* (pp.89-101). Addison Wesley Longman/ The British council.
- Savignon, S. J. (1997). *Communicative competence: Theory and classroom practice* (2nd ed.). Sydney: The McGraw Hill Companies.
- Saville-Troike, M. (2006). *Introducing second language acquisition*. Cambridge, UK; New York: Cambridge University Press.
- SBA Consultancy Team. (2005). *2007 HKCE English language Examination: Introduction to the school-based assessment component* (Training Package). Hong Kong SAR, China: Hong Kong Examination and Assessment Authority/Faculty of Education, The University of Hong Kong.
- Shohamy, E. (1983). The stability of oral proficiency assessment on the oral interview testing procedures. *Language Learning*, 33(4), 527-540.
- Shohamy, E. (1994). The role of language tests in the construction of second-language acquisition theories. In E. Tarone, S. Gass, & A. Cohen (Eds.), *Research methodology in second language acquisition* (pp.133-142). Hilldale, NJ: Lawrence Erlbaum.
- Shepard, L. (2000). *The role of classroom assessment in teaching and learning* (CSE Technical report 517). Los Angeles, CA: The Regents of the University of California.
- Stobart, G. (2006). The validity of formative assessment. In J. Gardener (Ed.), *Assessment and learning* (pp.133-146). Thousand Oaks, CA: Sage Publications.
- Thomas, R. M. (2003). *Blending qualitative & quantitative research methods in theses and dissertations*. Thousand Oaks, Calif.: Corwin Press.
- Thorndike, E. (1913). *Educational psychology Volume 1: The origin nature of man*. New York: Columbia University Teachers College.

Underhill, N. (1987). *Testing spoken language: A handbook of oral testing techniques*. Cambridge; New York: Cambridge University Press.

Vygotsky, L. (1978). *Mind in society: the development of higher psychological processes*. Cambridge, MA: Harvard University Press.

Woods, D., Bruner, J., & Ross, G. (1976). The role tutoring in problem solving. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 17, 89-100.

APPENDIX A: QUESTIONNAIRE

Section 1

1. I am a

- ① female
- ② male

2. How old are you?

- ① 23-30 years old ② 31-35 years old ③ 36-40 years old
- ④ 41-45 years old ⑤ above 45 years old

3. How many years have you been teaching in the secondary school?

중등학교 (중·고등학교 포함)에서 몇 년간 근무하셨습니다?

- ① less than 1 year ② 1-2 years ③ 2-5 years ④ more than 5 years

4. How many students do you have in a classroom? 교실당 학생 인원수

- ① 25-30 students ② 31-35 students ③ 36-40 students ④ 41-45 students
- ⑤ More than 45 students

Section 2

1. Do you include speaking assessment in your classroom? (Yes / No)

(현재 교실 말하기 평가를 실시하고 계십니까?)

→ If you answer 'No', please go to **No. 2**.

→ If you answer 'Yes', please go to **No. 3**.

2. What is the reason you do not assess learners' speaking competence in the classroom?

(교실 말하기 평가를 실시하고 있지 않다면 그 이유는 무엇입니까?)

- ① time constriction (시간부족)
- ② reliability issue (inappropriateness in high stakes test situation)
(신뢰도 문제; 고부담 시험환경에 부적합)
- ③ learners' insufficient proficiency (학습자의 유창도 부족)
- ④ annoyed about designing test battery including defining scoring criteria (평가 도구 제작의 어려움)
- ⑤ others: _____

3. The real purpose of assessment

Please indicate in effect for what purpose you employ classroom speaking assessment, and approximately how much times you typically spend on them in your curriculum.

현재 어떤 목적으로 교실 말하기 평가를 사용하고 계십니까? 평가에 소요되는 시간도 함께 표시해주시기 바랍니다.

	Time spent (hrs)
Assessment of students for class placement 수준별 반편성	0 1-2 3-5 5+
Pre-topic planning 학습 주제 선정	
Ongoing programming (lesson planning) 수업계획을 세우기 위해 학습 상황을 지속적으로 추적	
Ongoing student assessment (e.g. marking, feedback) 지속적인 누적 평가	
Final evaluation of topic/unit of work 한 교과과정이나 학습 단위에 대한 형성평가	
Providing information to others (e.g. bureaucratic report, parents, school) 교육관련 이해당사자들에게 정보제공	
Following the policy of schools or the education council 교육청이나 학교의 규정에 따르기 위한 목적	
Others:	

4. Please indicate the ideal purpose of classroom speaking assessment and the importance.

(most important = 5, somewhat important = 4, neutral = 3, less important, and least important = 1)

교실 말하기 평가의 이상적인 목적과 그 중요도를 표시해 주시기 바랍니다.

The ideal purpose of classroom speaking assessment	Importance
To place students in class 수준별 반 편성	1 2 3 4 5
To give learners feedback on progress 학습자에게 지속적인 피드백 제공	1 2 3 4 5
To give teachers feedback on learners progress 교사 자신에게 학습자의 발전과정에 대한 피드백 제공	1 2 3 4 5

To diagnose strengths and weaknesses and set further learning objectives 학습자의 강점과 취약점을 진단하여 학습 목표 설정	1 2 3 4 5
To provide information to others (e.g. bureaucratic report, parents, other teachers) 교육 이해 당사자 (학교, 학부모, 다른 교사들)에게 정보 제공	1 2 3 4 5
To indicate readiness for high school 고등학교 준비	1 2 3 4 5
To indicate levels of general speaking proficiency 일반적인 말하기 유창도를 알려줌	1 2 3 4 5
To indicate the students' achievement of course objectives 학습자의 수업 목표 성취도를 알려줌	1 2 3 4 5
Others 기타:	1 2 3 4 5

5. **Methods of assessment:** Please indicate 1) which of the following methods you use in your course; and 2) how many times you use it (or them) during a semester.

평가 방법: 1) 구체적으로 아래의 평가 도구 중 어떤 것을 사용하십니까? 2) 한 학기에 몇 번을 시행하고 계십니까?

<i>Methods of assessment</i>	Frequency
observation of students in typical speaking activities during regular classes 평상시 교실에서 이뤄지는 말하기 활동 관찰	
marking of a specific test tasks 말하기 활동을 점수화	
peer assessment 학습자간 동료평가	
self assessment 학습자 자기평가	
Others : 기타	

※ This is for those who choose the second option, 'marking of a specific test tasks' in No. 5.

6. **Test tasks:** please write *brief descriptions* of up to 4 test tasks/ activities which you employed during the past year for your classroom assessment of speaking and tick the given *student grouping*.

(e.g 'Role play: memorize the short dialogue given in the textbook and demonstrate in a

pair.')

현재 사용하고 계신 평가 도구가 학습자의 언어사용을 바탕으로 하는 경우 응답해주시기 바랍니다: 지난 일년 동안 채택하신 구체적인 말하기 활동 종류를 4가지만 간단히 기술하여 주십시오.

① (Student grouping: Individual / Pair / Small group / Whole class)

② (Student grouping: Individual / Pair / Small group / Whole class)

③ (Student grouping: Individual / Pair / Small group / Whole class)

④ (Student grouping: Individual / Pair / Small group / Whole class)

7. Please indicate the format of feedback that you use.

평가 후 피드백은 어떤 방식으로 제공되고 있습니까?

- ① only a score of marking
- ② criterion description plus score of marking
- ③ oral feedback given
- ④ written feedback given

Thank you for your participation. If you are willing to participate in the follow-up interview, please put a tick and let me know your contact details.

Yes, I can.

Contact details: (email) ***@***
(phone)

APPENDIX B: INTERVIEW QUESTIONS

■ *Background information*

1. How long have you worked in middle schools?
2. What is your language teaching approach?
3. What are assessment methods?
4. What is the language element focused on in the assessment?

■ *Teacher's perceptions of assessment effectiveness*

5. What do you perceive the function of classroom-speaking assessment?
6. Do you perceive any positive effects of classroom speaking assessment on teaching and learning?

■ *Difficulties of assessment administration*

7. If you have any difficulty in conducting classroom-speaking assessment, what are they?

APPENDIX C:

PARTICIPANT CONSENT FORM AND PARTICIPANT INFORMATION SHEET

PARTICIPANT CONSENT FORM (QUESTIONNAIRE)

I have read the accompanying letter explaining the project named “Current status of Classroom Speaking Assessment in Secondary Schools in South Korea”. The project is being conducted by Sujin Lee, a Master student of Applied Linguistics (TESOL), and supervised by Dr. Iwashita Noriko from the University of Queensland.

I understand that:

- The participation to this study is voluntary;
- I will be asked to participate in a questionnaire survey on classroom-based speaking assessment.
- It will take about 20 minutes;
- My confidentiality and privacy will be respected, future publications resulting from the study will use pseudonyms to ensure the anonymity of persons;
- I may request a copy of any publications arising from the work;
- I can withdraw my consent at any time without explanation. In that case, my record will be destroyed, and the data will be removed from the data analysis;
- I can contact Sujin.Lee on : sujin.lee@uqconnect.edu.au or +61 412 796 040 or her supervisor, Dr. Iwashita Noriko on n.iwashita@uq.edu.au to request further information about the project;

I _____ agree to participate in the above mentioned research project conducted by Sujin. Lee under the supervision of Dr. Noriko Iwashita from the University of Queensland.

Name _____ Signature _____ Date _____

PARTICIPANT INFORMATION SHEET

PROJECT TITLE:

Current practice of Classroom Speaking Assessment in Secondary Schools in South Korea

Dear teachers,

My name is Sujin Lee and I am conducting a research project under the supervision of Dr. Iwashita Noriko, a senior lecturer in the School of Languages and Comparative Cultural Studies, towards an Master of Applied Linguistics (TESOL) at the University of Queensland. Thus, I will be writing a thesis which is the equivalent of a 10,000 word report.

The overall aim of this study is to investigate the nature and quality of classroom-based speaking assessment currently administered in middle schools in Korea in terms of washback of assessment. As you may perceive, the absence of teaching and assessing speaking in the classroom of the formal school system in South Korea brings a gap in this specific research field. Thus, I hope that this study will contribute to providing secondary school teachers with more practical ideas of implementing speaking assessment in your classroom context. In addition to this, findings in the study will be able to promote educational authorities' perception of the need to support the assessment environment in the school context.

The study involves an online questionnaire as well as personal contact. If you consent to participate, you will be asked to fill out a questionnaire which consists of questions on demographic information and your methods of conducting classroom-based speaking assessment. It will take you approximately 20 minutes to complete the questionnaire.

Please note that the privacy of your responses will be respected. Your personal identification will remain confidential with your responses in any published accounts of this research. In the final report, you will be referred to under a pseudonym, not by your real name. The information you provide will be used for this research only and will not be passed on to any third parties. Your responses will be securely stored in a locked filing cabinet and will be destroyed through the secure disposal mechanisms of the University. A report of the study may be submitted for professional academic publication, but individual participants will not be identifiable in such a report.

Please also be advised that being in this study is completely voluntary and you are under no obligation to consent to participation. If you wish to withdraw at any stage, you are free to do so without prejudice simply by letting me know that you wish your information to be excluded.

There is a way through which you can consent to participate in this research. You can indicate that you have read and understood this information by signing the accompanying consent form. If you are interested in participating in the study or you require more information, you can contact Sujin, Lee on s4193855@student.uq.edu.au or +61 412 796 040.

Thank you.

Sujin, Lee

The study adheres to the Guidelines of the ethical review process of The University of Queensland. Whilst you are free to discuss your participation in this study with project staff (contactable on [insert your phone number here]), if you would like to speak to an officer of the University not involved in the study, you may contact the Ethics Officer of the School of Languages and Comparative Cultural Studies on (07) 3365 6311 (within Australia), +61 7 3365 6311 (international), or email at ethics@slccs.uq.edu.au

APPENDIX D: TEACHER PROFILES

	Variables	%	Frequency
Gender	Male	14	7
	Female	86	44
Age	23-30	12	6
	31-35	29	15
	36-40	29	15
	41-45	22	11
	above 45	8	4
Year of teaching	less than 1	2	1
	1-2	2	1
	More than 2-5	12	6
	more than 5 years	84	43
Class size (number of students)	25-30	8	4
	31-35	13	7
	36-40	43	22
	41-45	28	14
	more than 45	8	4
Conducting classroom speaking assessment	Yes	92	47
	No	8	4

APPENDIX E: DESCRIPTIONS OF SPEAKING TASK FORMATS

Category	Speaking task format	Task description	%	Frequency
		Role play	5.4	4
Responsive (40%)	(1) Role play	Role play in a pair: memorize a model dialogue and demonstrate it	16.22	12
		Role play in a small group: create a given situation and demonstrate it	2.7	2
	(2) Question and answer	Answer a set of questions given prior to the assessment	10.81	8
	(3) Mundane communication	Chat with “chatbots”	1.35	1
		Have a small talk with a teacher	4.05	3
Extensive (monologue) (40%)		Memorizing	1.35	1
	(4) Memorizing	Memorize English pop song	1.35	1
		Memorize useful sentences	13.51	10
	(5) Speech	Speech: demonstrate written script based on memorization	14.86	11
	(6) Story telling	Story telling	5.4	4
	(7) Description	Description	2.7	2
	(8) Reporting	Report what they acquired by independent web-based learning	1.35	1
	Interactive (13%)	(9) Interview	Interview: interviewing between a teacher and a student with preplanned interview questions	9.46
(10) Skit		Skit	2.7	2
(11) Pair work		Pair work	1.35	1
Intensive (4%)		Read aloud	1.35	1
	(12) Reading out	Read out student's own class reflection log in front of class	1.35	1
		Read a passage in the textbook	1.35	1
	(13) Vocabulary game	Explain vocabularies in limited time	1.35	1
Total			100	74

APPENDIX F: SUMMARY OF INTERVIEWEES' BACKGROUND INFORMATION

ID	Teaching experience (Year)	Teaching approach	Main assessment method	Focus of assessment
B	six	Mainly a grammar-translation method partially combined with a communicative language teaching method	Role-play with a dialogue given in a textbook and interview using suggested questions beforehand	Accuracy in the role-play task and fluency in the interview
H	ten	Both grammar-translation method and communicative language teaching method	Interview using a set of questions suggested beforehand	Fluency and accurate use of target grammar
O	seven	Mainly a grammar-translation method partially combined with a communicative language teaching method	Reciting useful sentences	None (i.e., giving points if students say something in English during the class)
S	eight	Mostly a grammar-translation method	Reciting useful sentences	Accuracy
Y	seven	Mainly a grammar-translation method partially combined with a communicative language teaching method	Question and answer: answer a set of questions given prior to the assessment	The extend of conveying meaning, especially knowing and using language features
P	six	Mainly a grammar-translation method partially combined with a communicative language teaching method	Speech: demonstrating a written script based on memorization	Fluency

APPENDIX G: INTERVIEW TRANSCRIPTION SAMPLE

Teacher Y's interview

1. How long have you worked in middle schools?

During my 7 year teaching experiences, I've been working in middle schools for 4 years.

2. What is your language teaching approach?

Now it is really difficult to implement communicative teaching approach in my context, while I think that more opportunities of practicing speaking through interaction is most important to improve oral competence. Also, I highlight the need to increase students' motivation, interest, and favor of learning English itself. Before teaching, it is first important to motivate Korean students' in learning English.

3. What is the language element focused on in the assessment?

I mainly focus on students' communicative ability. That is, even though they do not make a full sentence or show less grammaticality, I gave partial points to that if their answers make sense for successfully meaning delivery.

4. What are speaking assessment methods?

It is a question and answer type. Prior to the test, I suggest 20 questions which consist of various difficulty levels ranging low to high.

5. What do you perceive the function of classroom speaking assessment is?

6. Do you perceive positive effects of classroom speaking assessment on teaching and learning?

Now in my school, English communication class is particularly offered in addition to regular classes mostly focusing on reading comprehension and grammar instruction. So, I believe that classroom speaking assessment should be conducted in English communication classes, because it is in line with the objective of the course curriculum.

(Q: Are you now carrying out classroom-speaking assessment merely for the purpose of marking? Do you intend to use it for another purpose rather than marking?)

Yes. In fact, now I am carrying out classroom-speaking assessment merely for the purpose of marking, whereas if it is possible, what I mean is, some problematic conditions are solved, I'd like to use it to improve student's communicative competence. It may be at a very low level, though.

I do not think that assessment itself plays an important role to encourage students' speaking or to motivate their learning. In conclusion, if there is no compulsory instruction from the

educational ministry, I will not implement it to my classroom. My perception stems from the Korean context as an English acquisition, that is, English as a Foreign Language (EFL) context in which most students do not have opportunities to realize the need for English learning. In this situation, I do not want to give a lot of stress to students through assessment. Although I am forced to carry out classroom speaking assessment, I hope it will eventually facilitate student's knowledge or skills which are assessed. Moreover, as a teacher, I can identify the extent to which they have achieved the objective of the curriculum. However, they are not the purpose of my assessment, but I just hope it will occur as a consequence. I am doubtful about effectiveness of assessment. Students usually easily forget what they assessed on without enhancement. As I mentioned, they are given few opportunities to enhance what they have learned neither in the classroom nor in the real life.

7. If you have any difficulty in conducting classroom-speaking assessment, what are they?

I basically do not want to implement classroom speaking assessment in my teaching context at all. That's because, first of all, I always have too many students to assess. Second, classroom speaking assessment means that the given teacher who takes charge of the class should be an assessor. However, if marking scores reflect on the formal test, the issue of objectivity and reliability should be inevitably raised, because students belonging to the different teacher are assessed by a different assessor in spite of a common criteria sheet.

In addition, I do not have enough time to allocate for assessment. I am usually very busy to cover the curriculum within regular classes. Furthermore, I do not have time to teach speaking in class. Because I do not offer abundant opportunities involving speaking, I cannot insist on speaking assessment as well.

I heard that one teacher asked a native English teacher to solely evaluate classroom speaking assessment. Most students consent to these results without complaints, which is totally contrary to the situation in which she, the non-native teacher, played a role of assessor.