교육학석사학위논문

Korean English Teachers' Self-efficacy of Language Assessment Literacy and its Relation to Teacher Identities

한국 영어교사들의 평가 효능감과 교사 정체성의 관계

2022년 8월

서울대학교 대학원

외국어교육과 영어전공

이 수 민

Korean English Teachers' Self-efficacy of Language Assessment Literacy and its Relation to Teacher Identities

by

Soo-Min Lee

A Thesis Submitted to

the Department of Foreign Language Education

in Partial Fulfillment of the Requirements

for the Degree of Master of Arts in Education

At the

Graduate School of Seoul National University

August 2022

Korean English Teachers' Self-efficacy of Language Assessment Literacy and its Relation to Teacher Identities

한국 영어교사들의 평가효능감과 교사정체성의 관계

지도교수 소 영 순

이 논문을 교육학 석사 학위논문으로 제출함
2022년 6월

서울대학교 대학원
외국어교육과 영어전공
이 수 민

이수민의 석사학위논문을 인준함
2022년 7월

위 원 장 _____김 기 택_____

부위원장 _____김 현 우_____

위   원 _____소 영 순_____

Korean English Teachers' Self-efficacy of Language Assessment Literacy and its Relation to Teacher Identities

APPROVED BY THESIS COMMITTEE:

_____

KITAEK KIM, COMMITTEE CHAIR

_____

HYUNWOO KIM

_____

YOUNGSOON SO

ABSTRACT

Korean English Teachers' Self-efficacy of Language Assessment Literacy and its
Relation to Teacher Identities

Soo-Min Lee

English Major, Dept. of Foreign Language Education

The Graduate School of Seoul National University

Assessment plays an important role in effective language education. For
communicative language teaching in English education, the 2015 Revised
National English Curriculum emphasized the integration of teaching and
evaluation and expanded the implementation of process-oriented evaluation and
performance assessments. Accordingly, the importance of Korean English
teachers' assessment literacy as it affects effective assessment has come to the
forefront. However, several studies suggest that the assessment literacy of
English teachers in Korea is still low and the teachers have difficulties in its
development. To increase the self-efficacy of Korean English teachers in
language assessments, it is essential to scrutinize the influence of teacher identity
on the self-efficacy of teachers within the EFL context of Korean English

education.

The purpose of this study is to investigate the relationship between teachers′ self-efficacy in language assessments and the teacher identity. Exploring the relation between them will provide insights into improving the teachers′ language assessment literacy for more effective language assessments.

For this study, seventy Korean English teachers in secondary schools voluntarily participated in an online survey. The online questionnaire consisted of 49 items and was divided into two subsections: Section A regarding the self-efficacy of language assessments, and Section B about the teacher's identity as a professional. Then, nine teachers, who voluntarily participated, were randomly chosen for subsequent individual interviews for further investigation. They shared their own stories on their self-perceived efficacy of language assessments and their self-identification as a professional language teacher. The interviews were thematically analyzed and the themes found from the analysis were categorized into two constituents of the identity formation: discourse and practice.

The findings suggested that a significant correlation existed between the self-efficacy of teachers and the positive teacher identity of Korean English teachers. The teachers who identified themselves as experts showed high self-efficacy in their assessments. Moreover, it suggested that self-awareness as a

non-native English speaking teacher in EFL circumstances affects both teachers′ assessment efficacy and the formation of their identity as experts. In terms of discourse, they had identity conflict as they possessed both learner identities as non-native speakers and teacher identities. With regard to practice, the teachers seemed to position themselves as experts by valuing teacher collegiality with strong teacher agency. However, their expert identity still seemed to be at risk because the exonormative norms, which undermine Korean English teachers′ expert identity, are still prevailing in their assessment practices.

Although the study had some limitations in terms of methodological aspects and the small number of study subjects, it will contribute to a better understanding of Korean English teachers′ self-efficacy in their assessments. Moreover, by investigating its relation to their identity construction, the study shed light on practical suggestions that can support teachers to develop the language assessment literacy and improve their assessment practices.

Key Words: teacher self-efficacy, assessment efficacy, teacher identity, teacher efficacy, language assessment literacy, foreign language assessment, Korean English teachers

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1.

INTRODUCTION

This study aims to investigate Korean English teachers' self-efficacy in language assessment and its relation to their teacher identity. This chapter outlines the purpose and organization of the thesis. Section 1.1 introduces the motivation and purpose of the study, followed by research questions in Section 1.2. Section 1.3 describes the organization of the thesis.

1.1.   The Motivation and Purpose of the Study

The movement toward Communicative Language Teaching (CLT) ushered in a new era in the field of language learning. The distinction between competence and performance made educators aware that if language was to be learned for its true communicative purposes, then the classroom tasks must also reflect the communicative and performance aspect of language behavior. As such, language tasks and assessments began to move beyond the purposes of reviewing learners' intrinsic linguistic competence, to include how well learners are able to demonstrate linguistic knowledge and

skills within communicative contexts (Black & Wiliam, 1998; McNamara & Roever, 2006). This new attention to performance has resulted in in-class practices and assessments with communicative learning goals that emphasize process over product. Many language teachers, however, have experienced difficulty in creating such tasks and evaluations. The burden becomes enhanced for Korean English teachers(KETs) as well, as they are often less familiar with these types of assessments than traditional paper-based assessments focused on language competence and they may question their ability to assess learners' language behavior (Park & Chang, 2017).

Accordingly, the development of English as a foreign language (EFL) teachers' competency in language assessment, or language assessment literacy (LAL), has come to the forefront (Fulcher, 2012; Taylor, 2013). The teachers, since they work as both language teachers and test makers, need to acquire a high level of language assessment literacy to fulfill their responsibility (Taylor, 2013). Given that the Korean 2015 Revised National English Curriculum (RNEC) emphasizes process-oriented assessment and the integration of teaching, learning and assessment to promote student-centered learning, English teachers in Korea are required to administer alternative forms of assessments including performance-based process evaluations as well as traditional paper-based tests. To do so, they employ performance assessments such as presentations, interviews, or essay writing in order to

assess their students' language ability (Lee, 2018; Park, 2016). Indeed, their ability to understand and engage learners through a range of assessments has become essential.

Among the diverse stakeholders, classroom English teachers can be accounted as the "real protagonists" of assessment, as they deal with external and internal assessment procedures on a daily basis and are closer to the heart of the overall assessment procedures (Erickson, 2020, p.40). The teacher, hence, is a crucial factor in effective classroom assessments (Vogt & Tsagari, 2014). However, while teachers' assessment practice has been emphasized, their LAL competency has been questioned and studies also suggest that teachers' LAL seems to be underdeveloped (Hasselgreen, Carlsen & Helness, 2004; Jeon & Oh, 2006; Vogt & Tsagari, 2014). Korean English teachers are also reported as experiencing difficulties in assessing students especially in retaining reliability, objectivity, and practicality of the test (Park & Chang, 2017). To provide practical support for the LAL development of teachers, identifying the root of a discrepancy between teachers' LAL knowledge and their assessment practice is needed.

Existing studies of LAL focused on the constructs of LAL by identifying the elements of language assessment knowledge and skills that language test stakeholders, including classroom language teachers, need (Fulcher, 2012; Hasselgreen et al., 2004; Stiggins, 1991). Also, research has

pinpointed incomplete pre-service or in-service training as a major reason for the underdevelopment of teachers' LAL and investigated teachers' training experience, their current LAL state, and training needs (Chung & Nam, 2018; Fulcher, 2012; Hasselgreen et al., 2004; Vogt & Tsagari, 2014). Less attention, however, has been given to primary sources that aggravate their negative self-image as disqualified assessors and keep KETs away from pursuing self-reflective development of LAL in their practice.

Teachers are under the pressure of performing dual roles as teacher and assessor and, yet, they are also expected to be self-accountable. Moreover, they are required to deal with the recent alteration of language education theories which focuses on communicative and sociocultural aspects of language. This modified paradigm contrasts with former theories which worked on the accumulation of linguistic knowledge. The paradigm shift of language learning compels teachers to incorporate sociocultural theories of language learning in their assessments. Hence, various types of language evaluations including performance assessments or alternative forms of assessments are developed and suggested for language teachers through pre-service and in-service training. Even so, the teachers still think their own LAL level is insufficient (Chung & Nam, 2018; Jeon & Oh, 2006; Park & Chang, 2017; Vogt & Tsagari, 2014). To develop teachers' LAL and their assessment practices, the complex framework of assessment, which is

influenced by the conceptual transition of language learning between contrasting paradigms, needs to be considered as well as providing an expanded repertoire of assessment techniques (Scarino, 2013). Understanding how teachers relate their practices to their framework of language learning and assessments is essential for improving language assessment praxis. Not only should teachers be capable of understanding the defined knowledge base of language assessment, but they must also be aware of it and relate their knowledge to practice in their particular classroom context.

Very few studies have investigated the LAL of Korean English teachers (KETs) and existing studies of KETs' LAL have focused only on their needs for teacher training based on the knowledge base (Chung & Nam, 2018; Jeon & Oh, 2006). According to Ha and Min (2008), the development of teachers as professionals requires an internal change of attitude and identity, while training generally includes learning skills only; therefore, aspects other than lack of teacher training need to be investigated in order to understand the underdevelopment of teachers LAL. In this regard, existing studies have limitations as they examined LAL as skills that can be learned through training and applied by teachers without considering the impact of social context, especially in the EFL context of language assessment in Korea.

The socio-cultural context of English education in Korea strongly influences the development of the teacher competence of KETs. The native-

nonnative distinction in the EFL context influences teacher identity construction which is closely related to the development of teacher competence (Aneja, 2016). Researchers have noted that some Korean English teachers identified themselves as non-native speakers and considered their language features as a source of vulnerability for their teaching. Especially, they felt less qualified in their speaking ability as they are ashamed of their accents and assumed having native-like pronunciation is one of the key requirements for a qualified English teacher. Based on those assumptions, they viewed themselves as less qualified teachers than native English speaking teachers (NESTs) (Choe, 2008; Hong, 2013). The native-nonnative divide bolsters the discourse of the native speaker model in Korea, which disempowers Korean English teachers as non-native English speaking teachers (NNESTs).

Moreover, some researchers found that English ability is considered as not only foreign language ability but also cultural capital, which symbolizes a privilege of the elite in Korea (Choe, 2008; Hong, 2013; Kim, 2013; Yi & Yang, 2009). Based on this notion of 'English is a power', some Korean English learners prioritize having native-like fluency and pronunciation even from the pre-school level of language learning and preferred NESTs over NNESTs (Hong, 2013; Yi & Yang, 2009). Based on the prevailing native-speaker model, KETs are continuously compared to NESTs

in terms of language proficiency or teaching skills and have been reported showing lower self-efficacy in teaching compared to teachers of other subjects such as Korean literature or math (Kwon, 2013). Based on the native speaker fallacy, which idealizes the native speaker model, KETs deemed themselves less qualified in certain language skills. For instance, some KETs showed a high level of anxiety in teaching speaking skills and even tried to avoid speaking in English in their classes. They thought that the anxiety comes from NNEST status and can be overcome by having some experience in English-speaking countries (Choe, 2008; Hong, 2013).

Within the discourse of the native speaker model, NNESTs struggle to negotiate their professional identity (Choe, 2008; Aneja, 2016). To develop KETs' LAL and improve language assessment practice, the contextual aspect of KET's LAL should be investigated. Scarino (2013) reported developing teachers' LAL necessitates the interplay of the theoretical knowledge base and teachers' interpretive frameworks. The frameworks of language assessments are shaped through teachers' personal experiences, beliefs, and understanding of language education and language itself, and are highly affected by their identities. Therefore, in order to understand the status quo of KETs' LAL, the influence of NNEST identity on LAL development needs to be investigated.

1.2. Research Question

The present study is aimed at identifying KETs' self-efficacy of LAL in assessing language skills and its relation to teacher identity. In this study, the teachers' self-efficacy of LAL will be investigated in detail to determine how and to what extent teacher self-efficacy of LAL differentiates based on the domains of the four language skills and whether the efficacy relates to their NNEST identity. It is guided by the following research questions:

1. What is Korean English teachers' self-efficacy of their LAL in each language skill?

2. How does the NNEST identity relate to the teachers' self-efficacy of LAL?

1.3. Organization of the Thesis

This thesis consists of six chapters. Following this introduction chapter, Chapter 2 presents a review of the literature on language assessment literacy and teacher identity that helped to form the research question of this thesis. Chapter 3 describes the research methods including instrument, participants, and procedures of the study, followed by the statistical and qualitative measures taken for the data analysis. Chapter 4 reports descriptive

statistics of teacher self-efficacy of LAL and their expertise identity and discusses findings of relations between self-efficacy and identity formation. Chapter 5 discusses the noteworthy issues that arose from the results with regard to the research questions. Lastly, Chapter 6 summarizes the research findings with pedagogical implications and concludes with the research limitations and the future research suggestions.

CHAPTER 2.

LITERATURE REVIEW

2.1 Language Assessment Literacy (LAL)

2.1.1 New Language Assessment Practices

Rising of the communicative language teaching (CLT) caused a shift in understanding language proficiency and language assessment. Language proficiency was previously defined as knowledge of structures in phonological, grammatical, and lexical aspects of language. In the era of CLT, however, integrating language knowledge in performance was emphasized more than knowing language structures, as using language in communicative way was considered more significant as evidence of one's linguistic ability. Hence, a renewed focus was given to communicative testing including performance tests and other alternative forms of testing, away from norm-referenced tests with discrete-point items (Black & Wiliam, 1998; McNamara & Roever, 2006, p.43).

In classroom language learning, the role of assessment has been expanded to include not only measuring students' current competence but

also diagnosing their developmental needs to promote learning itself (Black & Wiliam, 1998). Language assessments aimed to infer students' current language abilities from their language behavior. Many assessment theories and methods were hence developed to make language assessment more systemic and principled. Off-the-shelf assessments, however, were unsatisfactory for facilitating students' learning and assessing their linguistic progress. Moreover, with the advent of communicative language learning, a change in language assessment was required. Consequently, classroom-based assessments were emphasized with an increasing awareness of their direct impact on language learning (Black & Wiliam, 1998; Leung, 2007; Hill & McNamara, 2012; Rea-Dickins, 2001, 2006). The interactions between classroom assessment and learning have been investigated extensively, and the importance of assessment for learning has been illuminated by the insights from the investigation (Black, Harrison, Lee, Marshall & Wiliam, 2004).

To embrace this change in the assessment framework, the 2015 Revised National English Curriculum (RNEC) in Korea draws attention to integrating assessment into learning to promote student-centered learning. Accordingly, classroom English teachers in Korea are employing different types of assessments including performance tests, essay writing, or student portfolios. The majority of the teachers are aware of the educational benefits

of communicative assessments in terms of the validity of the test and the positive effect on students' learning (Kang, Jin & Jang, 2015; Park & Chang, 2017). However, researchers found that the teachers are facing difficulties in assessment practices related to validity, practicality, and reliability issues (Bae & Shin, 2016; Bang & Chon, 2011; Park & Chang, 2017). They experienced difficulties in establishing standards and faced troubles in assessing students speaking directly. Moreover, external forces such as national policy regulations on school assessments and a high priority on college admission test preparation strongly restrict teachers' assessment practice (Park & Chang, 2017; Sung & Jo, 2015). To handle the complicated dynamics of classroom language assessments in the best possible way, therefore, teachers need to develop their LAL which requires understanding their educational context.

2.1.2 Defining Language Assessment Literacy (LAL)

In theoretical and empirical research, language assessment literacy (LAL) and more general assessment literacy (AL) are relatively new fields and LAL is considered an overlapping or subordinate category of AL (Taylor, 2013). The concept of assessment literacy encompasses the knowledge and skills which are needed to evaluate students' abilities. LAL, on the contrary,

was defined as the integration of language-specific competencies into layers of assessment literacy (Inbar-Lourie, 2008). LAL is a multifaceted concept and definitions of LAL have differed within assessment literacy literature. Early contributions of LAL research focused on defining LAL components and LAL profiles for stakeholders (Davies, 2008; Fulcher, 2012; Inbar-Lourie, 2008; Jeong, 2013; Malone, 2013; Pill & Harding, 2013; Taylor, 2009). Still, this tendency is prominent in recent studies, in terms of both LAL and AL.

LAL research established the knowledge base which is needed to conduct language assessment procedures. It includes every phase of assessment such as planning, administering, and interpreting language assessment data, and recently, with the constructivist view, socio-cultural aspects of assessments are added to the initial knowledge base of LAL (Fulcher, 2012; Inbar-Lourie, 2008). LAL includes 'what', 'why', and 'how' of the language assessment. Therefore, the trait and methods of language testing and assessments are included in LAL components as well as background reasoning of the socio-cultural aspects of language assessments.

Narrow definitions of AL suggested a priori standards of LAL to be learned by language teachers. For example, Popham (2011) presented that AL is composed of what an individual understands about assessment concepts and procedures which influence one's educational decisions. Those narrow definitions contributed to the tendency of AL research as the research draw

heavily upon theory and technical issues of test validity, rather than empirical research. LAL research has avoided such predetermined AL standards and attempted to discover which aspects of language and language assessments should be taken into account for defining and developing LAL (Erickson, 2020).

Researchers also focused on deriving a working definition of LAL and to do so, they tried to reveal the LAL needs of stakeholders. Fulcher (2012) conducted a survey to discover English teachers' LAL needs and carried out factor analysis assisted by qualitative analysis of participants' interviews. Then, he proposed a working definition of LAL as follows:

The knowledge, skills, and abilities required to design, develop, maintain or evaluate, large-scale standardized and/or classroom-based tests, familiarity with test processes, and awareness of principles and concepts that guide and underpin practice, including ethics and codes of practice. The ability to place knowledge, skills, processes, principles, and concepts within wider historical, social, political, and philosophical frameworks in order to understand why practices have arisen as they have, and to evaluate the role and impact of testing on society, institutions, and individuals. (Fulcher, 2012, p.125)

Fulcher claimed that language teachers are already aware of various assessment demands. So he said that LAL has to be discussed within a wider social context to meet teachers' needs. He made an attempt to connect practice, theory, and context by providing a cyclic test designing process. According to his working definition of LAL, decision-making in assessment practice is informed by theory, and the theoretical exploration is realized in the practice of language testing. Fulcher's extended dimension of LAL is illustrated in Figure 2.1.
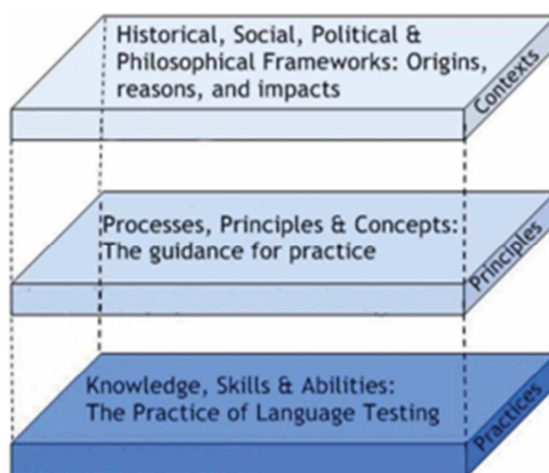
Figure 2.1. Language assessment literacy: an expanded definition (Fulcher, 2012, p.126)

Even though he noted that these components are not prerequisites for all stakeholders, Fulcher's working definition of LAL highlights the importance of a larger social and contextual framework for language assessments. LAL definition of knowledge base involves understanding and

practicing assessment properly with theoretical knowledge and adequate skills (DeLuca & Klinger, 2010). Expanded LAL definition involves contextual and consequential validity and restructures LAL in micro and macro dimensions. In this regard, in developing teachers' LAL, it is essential to develop teachers' ability to self-evaluate their preconceptions of language and assessment so they can interpret their own assessment practices and learner's language behaviors with contextual considerations. With this developed self-awareness as assessors, teachers can better understand the assessment phenomena and ultimately transform assessment practices (Scarino, 2013).

Taylor (2013) suggested eight components of LAL: knowledge of theory, scores and decision making, principles and concepts, local practices, technical skills, language pedagogy, personal beliefs/attitudes, and sociocultural value. The research presented different LAL profiles of test constituencies and illustrated different dimensions and components of LAL are required for different groups of stakeholders, due to the contextual diversity of assessment. Figure 2.2 shows a LAL profile of classroom English teachers and it is suggested by Taylor (2013).
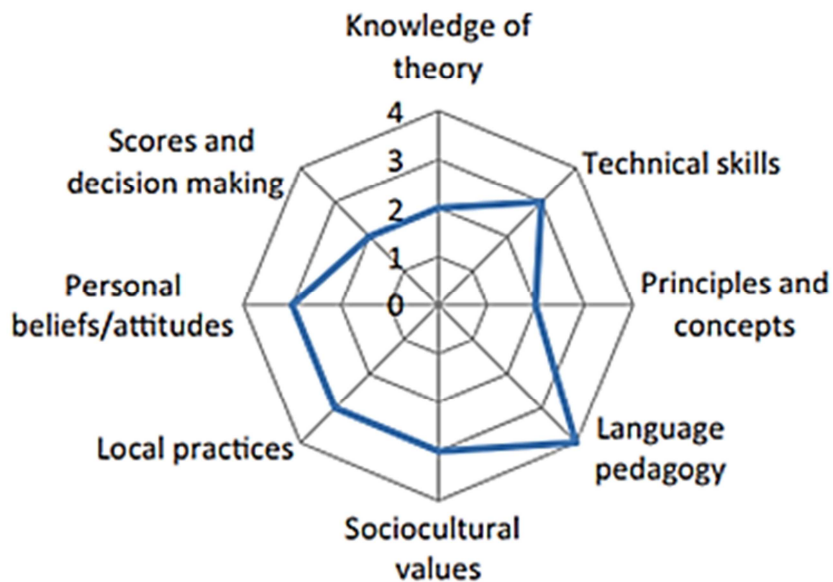
Figure 2.2. Classroom teachers' LAL profile (Taylor, 2013, p.410)

The diagram is based on hypothesized dimensions of possible LAL components from various LAL research. It shows that elements of LAL include multidimensional aspects of language assessment such as sociocultural values, and personal beliefs and attitudes. Those contextual and personal aspects would also compose individuals' interpretative framework of assessments, which needs to be discussed for developing one's LAL (Scarino, 2013). In this regard, recent studies of language education clarify that LAL is not a decontextualized concept. Hence, being aware of the sociocultural context of language assessment is needed as a precursor to developing teachers' LAL and assessment practice.

2.1.3　LAL Training for EFL Teachers

EFL teachers' LAL has received much interest in TESOL research, due to the phenomenal increase of assessment responsibilities on language teachers accompanied by expanded use of language assessment results across various social domains. Many research projects aimed to measure the language teachers' current practice of language assessment and elicit their LAL training needs with surveys based on theoretical backgrounds of assessment (Fulcher, 2012; Hasselgreen et al., 2004; Vogt & Tsagari, 2014). The research made contributions to providing new content for teacher LAL training by uncovering the changing needs of language teachers for a new age.

Fulcher (2012) designed a survey for language teachers based on four domains: designing and developing tests, classroom assessments and washback, large-scale standardized testing, and validity and reliability. The survey was delivered globally over the Internet and results were analyzed with factor analysis. The research revealed the desire of teachers to understand more assessment principles as well as the practical "how-to" of assessments.

Hasselgreen et al. (2004) also surveyed to unveil the LAL training needs of European teachers, teacher trainers, and experts. According to the survey, EFL teachers in Europe rarely had the pre-service LAL training and

felt they need to be trained especially for informal assessments and practical skills of developing tests. Different stakeholders had different needs of LAL but every stakeholder group coincided with the need for formal education and training in LAL. The research also reported that in developing teachers' LAL, the most significant areas are related to assessment for learning and formative assessments: portfolio assessment, feedback, pear- and self-rating assessments, rating performance tests, and so on.

Vogt and Tsagari (2014) aimed to reveal the gap between the kinds of LAL training EFL teachers in Europe have received and the LAL training they need. The survey explored three components of LAL: purposes of testing, classroom-focused language testing and assessments (LTA), and LTA concepts and content. The result of the study indicated that European teachers' language testing and assessment literacy is underdeveloped and most teachers resorted to strategies to compensate for their lack of LAL. They have learned on the job and get help from their colleagues and textbook materials when they make LTA decisions. Also, pre-service or in-service training was not sufficient for the practical assessments. The research findings somewhat corroborate the study of Hasselgreen et al. (2004) in that both found teachers' strong need for LAL training. On the other hand, it differed in the areas of perceived developmental needs. While the former research report highlighted that an apparent needs for teacher training was found on alternative forms of

tests, this research presented teachers also wanted further LAL training across the range of LTA features including more conventional areas, such as the assessment of reading and grammar as well as alternative forms of assessments like a portfolio. Moreover, Vogt and Tsagari also revealed that teachers strongly required training for criterion-oriented assessment of speaking skills as they perceive their own oral assessment practices as holistic and subjective.

In Korea, much research attempted to reveal KETs' current LAL and their training needs through surveys to provide adequately designed teacher LAL training sessions. The majority of the studies, however, were focused on curriculum and general assessment-related factors rather than including linguistic specificities or intercultural aspects of language assessments.

Jeon and Oh (2006) revealed a discrepancy between Korean English teachers' self-evaluated competence of LAL and their perception of the importance for each component of LAL in secondary schools. Five components of LAL were surveyed: knowing properties to be assessed, utilizing teaching procedure and assessment, developing the assessment method, understanding the quality of the assessment, and producing valid and reliable assessment results. The research shows that the teachers think their own LAL was deficient in the areas in which they have not been trained, such as evaluating the reliability and validity of an assessment, understanding

characteristics of certain language behavior to be tested, and planning a specific test for the language behavior. More importantly, according to the research, KETs perceived themselves as their LAL competence has not reached a level of expertise in all components of LAL compared to their high awareness of the importance of those LAL components.

KETs' previous LAL training experiences and needs were also investigated (Chung & Nam, 2018). A survey and follow-up interviews were conducted to measure KETs' experiences and further need in LAL development. The research reported that KETs were not satisfied with their previous LAL training and expressed a strong need for training in overall LAL areas, especially in the areas of planning performance assessment and testing productive skills. With regard to assessment practice, the results presented that KETs use some strategies to compensate for their lack of LAL such as modeling test items or seeking advice from senior or co-working teachers to deal with insufficient LAL competence and the research suggested that current LAL training may not meet the practical needs of KETs.

Thus, the existing literature is mostly focused on examining teachers' current assessment practice and their training needs to support KETs' LAL development by providing well-organized LAL training. Considering the definition of LAL in recent studies, investigating KETs' LAL concerning its multidimensional and contextual aspects is essential as a precursor to

understanding and developing KETs' LAL.

2.2   Non-native English Speaking Teacher (NNEST) Identity

Considering the worldwide use of English and unprecedented global influence of it, English gains the status of English as English as a Lingua Franca (ELF), or an international language (EIL). It is estimated that the number of ESL or EFL speakers outnumbers native speakers or the speakers of inner-circle varieties of English (e.g. American or British English). Braine (2010) commented that in terms of worldwide English education, approximately 80% of English teachers are non-native speakers of English. As the notion that English should be taught as an international language grows, the dominating norm of 'Standard' English based on the ideology of the native-speaker model is challenged in English language teaching (ELT) (Holliday, 2005).

Despite the imprecision of the terms *nativeness* and *nonnativeness*, they construct the "bedrock of transnationalized ELT" (Leung, 2005, p.128). This dichotomy influences many aspects of ELT such as proficiency tests (American Council on the Teaching of Foreign Languages, 2012), the model of language use (Nguyen, 2017), as well as teacher identity and teacher development (Aneja, 2016; Choe, 2008). As Phillipson (1992) stated the

concept of a native English speaking teacher (NEST) was constructed as a superior and even ideal English teacher, due to the paradigm of linguistic imperialism. According to the ideology of native-speakerism, NESTs stand for a "Western culture" that originates the English language ideal and its teaching methodology (Holliday, 2006, p.385). Holliday (2005) also noted that native-speakerism implies an assumption that native-speakers of the English-speaking West possess a special claim for the language itself, thus it secures the superiority and authority of the native speakers.

A reified notion of idealized NEST includes that people from English-speaking West can only be perceived as native English speakers, the only source of "correct" English. As a result of this implicit reflection of linguistic imperialism, NEST status is selectively given to native speakers who speak certain English language varieties, and teachers without such status are continuously questioned in their professional competency of teaching (Aneja, 2016). Aneja (2016) mentioned that in some cases, even native speakers who speak English other than 'standard' English (e.g. American or British English), such as speakers of African American Vernacular or Indian accented English, are questioned in their legitimacy of NEST in the TESOL profession even though their native language is English. Therefore, within the continuing hegemonic influence of native-speakerism,

nonnative English speaking teachers are disempowered and it has been hard for them to establish themselves as qualified language teachers or assessors.

The new paradigm of EIL rejects such divisive ideology of native-speakerism and raises issues in multi-competence perspective to challenge the native-speaker norm (Cook, 2016). In the NNEST movement, studies have focused on the strengths of NNESTs such as having experience as L2 learners themselves and sharing the same cultural and linguistic backgrounds with their students (Medgyes, 1992). Hence, the NNEST's frame as the perpetual learner is problematized (Brutt-Griffler & Samimy, 2001) and a recent study suggests the notion of native or nonnative speakers should be reconceptualized as negotiated social subjectivities which are multiple and dynamic, rather than distinguishing NNESTs from NEST and advocating NNESTs' strength (Aneja, 2016).

In recent TESOL studies, the idealized notion of NEST is continuously questioned (Canagarajah, 2013). Yet, the native speaker fallacy (Phillipson, 1992) seems to be still effective in periphery communities which include Expanding or Outer Circle countries such as East Asian countries like Korea. Researchers reported that English teachers from the periphery still express exonormative tendency, favoring native-speakerism, with negative attitudes towards non-native English varieties and English learners tend to believe near-native accents based on General American or Received

Pronunciation are acceptable (Jenkins, 2009; Monfared & Khatib, 2018). Nguyen (2017) found that TESOL teachers in Vietnam who are already aware of and appreciate different English varieties still identify strongly to the native-speaker model when it comes to their own language proficiency. Especially in terms of pronunciation and accents, they worried about the non-nativelike features of their language use and fall back into the native-speaker fallacy.

English education in Korea also seems to hold persisting stigmatization based on the native-nonnative divide. Some researchers pointed out that native-speakerism is embedded subconsciously in the mindset of Korean English teachers, parents, students, and school administrators (Bae, 2015; Choe, 2008; Hong, 2013; Shin & Park, 2013). A socially constructed notion that 'English is the ability and power' plays a significant role in the teachers' identity formation (Bae, 2015; Hong, 2013; Kim, 2013). Similar to NNESTs in the ESL context (Aneja, 2016), native-like pronunciation is highly valued in judging the English oral proficiency of KETs, and teachers with L1 accents were seen as less qualified and showed less confidence in their classroom performance (Choe, 2008; Hong, 2013).

Within the dichotomous discourse of English, KETs struggle to establish a professional teacher identity. Some of them try to reconstruct positive identity through acquiring additional TESOL certificates from

English-speaking countries or calling on cultural experience in those countries (Choe, 2008). Some Korean English teachers are aware of their NNEST status and suffered from anxiety about their lack of English proficiency, regardless of their proper command of English and teaching ability. For instance, Choe (2008) reported that the teachers tend to be preoccupied with the thought that NNESTs' linguistic mistakes are likely to be interpreted as a lack of linguistic proficiency. According to Hong (2013), due to this anxiety, KETs framed themselves as perpetual English learner and felt continuous tension to fulfill the expected language proficiency level. The negative perceptions of NNEST influence the construction of Korean English teacher identity significantly and hinder them from identifying themselves as professional experts.

Teacher identity is dynamic, multiple, and shifting as it is repeatedly negotiated and recreated within contextual influences (Aneja, 2016; Pennington, 2014). Trent (2015) proposed a multifaceted and multidimensional framework for empirical research of teacher identity. The integrated framework for teacher identity investigation reflects identity the construction in both discursive and experiential ways. According to the framework, both "identity-in-discourse" and "identity-in-practice" are essential constituents of teacher identity (Trent, 2015, p.46). "Identity-in-discourse" describes identity as it is formed and negotiated through language.

Individuals locate themselves within a discourse and identify themselves by overt naming and making commitments to those names of identities. "Identity-in-practice" indicates that the identity formation process is operationalized through concrete practices that people experience, and individuals also actively identify themselves by behaving in certain ways (Varghese, Morgan, Johnston & Johnson, 2005). Trent suggested that both of these approaches require attention in the exploration of identity and reflected them in the framework of identity formation. Figure 2.3 summarizes the framework of identity formation and illustrates the role of language and practice in it.
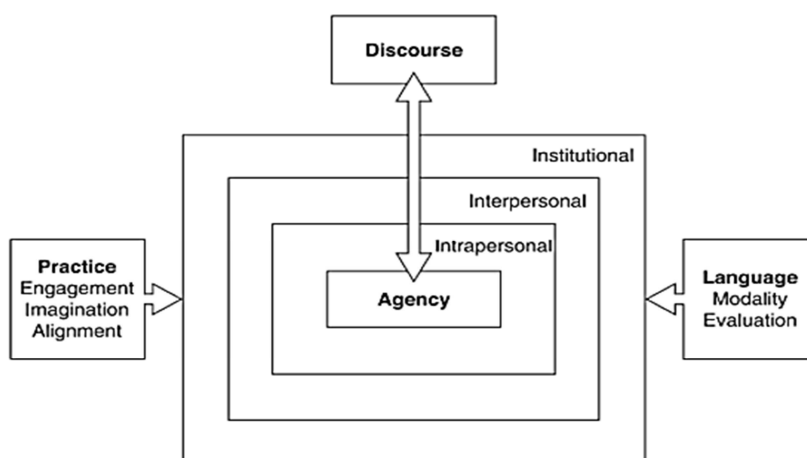
Figure 2.3. A framework for teacher identity investigation (Trent, 2015, p.47)

The framework suggests that identity reflects the influence of language and practice on multiple levels. In the framework, agency and discourse are illustrated as they affect each other in the process of identity

formation. Individuals gain subjective positions from the discourse and interpret the world around them actively within the discourse, while the discourse reflects individuals' attitudes, beliefs, and values.

In the framework of identity construction, language and practice represent identity-in-discourse and identity-in-practice respectively. The discourse is manifested through language, which consists of modality and evaluation. Modality is defined as what is committed as truth and obligation, and evaluation indicates what is considered to be desirable and both invoke implicit value systems of identity construction. Individuals discursively locate themselves by the overt naming of their own identities. The experiential construction of identity is conceptualized in three different modes: engagement, imagination, and alignment. According to the modes, individuals engage in relations with others to acquire the conventions of the discourse community, gain membership in imagined communities, and coordinate their activities within a broader context of organizations. The framework suggests that significant considerations are needed both in discursive and experiential ways of identity construction to reveal the ways in which structure and agency interact. Hence, a comprehensive exploration of NNEST identity requires consideration of the positionality of NNEST within the discourse and practice of English language education.

Existing literature, indeed, reported that the marginalized positionality of NNEST can enact alternative teacher identification as a multi-linguistic and multicultural speaker or as a specialist in L2 teaching, rather than the language itself in both EFL/ESL contexts (Aneja, 2016; Canagarajah, 2004; Hong, 2013; Jenkins, 2009; Monfared & Khatib, 2018; Nguyen, 2017). Teacher identity formation is reported to be highly pertinent to the development of teacher professionality (Kim & Cheong, 2012; Lee, 2010).

Hence the complex and social nature of negotiated NNEST subjectivity needs to be considered a significant factor in the studies of teacher LAL. Contextual influence on teacher LAL is implied in some LAL research in EFL or ESL context. For instance, in the EFL context, NNEST teachers have shown a tendency to feel least qualified in assessing students' productive skills than in any other area (Hasselgreen et al., 2004). Kang and Lee (2012) found that KETs possess a strong tendency to follow the native-speaker model and withhold accepting EIL features when they are correcting learners' errors. Therefore, accordingly, to understand KETs' assessment experience and their self-efficacy of LAL, the research focus should be reoriented to identity-driven approaches and it is essential to include NNEST identity in the focus of research.

2.3 Limitations of Previous Research

Much research has been conducted on LAL and has revealed many significant factors needed for developing LAL of language teachers. However, considering the macro and micro dimensions of LAL suggested by Fulcher (2012) and Taylor (2013), previous studies investigated the concept of LAL in limited dimensions of the knowledge base (Chung & Nam, 2018; Hasselgreen et al., 2004; Jeon & Oh, 2006; Vogt & Tsagari, 2014). In this regard, sociocultural dimensions of LAL are less investigated. Taking into account of macro dimension of LAL, it is necessary to consider the connection between the LAL implementation and teacher identity in the theoretical and social contexts.

The components of LAL regarding teachers' personal beliefs and attitudes, and sociocultural values overlap with issues of teacher identity construction. Scarino (2013) argued that teachers' preconception, beliefs, and personal experiences construct their interpretive framework, and raising awareness of the framework is essential for developing teacher assessment literacy and transforming assessment practice. Despite the breadth of existing literature, there is a lack of exploration on the teacher identity and its relation to LAL, which is needed for understanding the teachers' perception and implementation of LAL.

Moreover, the self-efficacy of teacher expertise and teacher identity is found to be interrelated (Kim & Cheong, 2012; Lee, 2010). Few empirical studies, however, have explored an application of identity-driven approaches to LAL in English teachers in Korea. Hence, the relation between NNEST identity and LAL is a noteworthy aspect for understanding the LAL of Korean English teachers and needs to be researched further. In the light of these considerations, this study aims to explore KETs' self-efficacy of LAL, its relation to NNEST identity, and the contextual and sociocultural sources that influence their LAL efficacy beliefs.

CHAPTER 3.

METHODOLOGY

This chapter contains clarification of the methodological approach and research design in the study. This study adopted a mixed-method approach to examine research questions. A mixed methods approach can be defined as a combination of qualitative and quantitative methods in single research (Dornyei, 2007). Through this combined analysis, quantitative and qualitative inquiry can interpenetrate and inform each other. Since this study aims to investigate the complex relation between KETs' LAL self-efficacy and their teacher identity, analyzing the numeric tendency of these factors and understanding specific societal contexts which add meanings to numbers are both needed. Hence, a multi-level analysis of quantitative and qualitative data was chosen for the study. Section 3.1 depicts specific information about participants, and section 3.2 explains instruments employed in this study. Lastly, section 3.3 describes the general research procedure of data collection and analysis.

3.1 Participants

A total of 70 Korean English teachers in secondary schools were recruited for the research survey. A posting requesting participation in the survey was posted on an online community forum of nationwide in-service English teachers. The posting fully explained the purpose and procedure of the research, and all participants voluntarily applied to enroll in the research. English teachers who are Korean and worked at a middle or high school were eligible for participation. At the end of the questionnaire, there was an item asking for consent to participate in a following semi-structured interview after the survey. Among those who voluntarily agreed to take part in the follow-up interview, a total of nine interviewees were randomly chosen.

3.2 Instruments

3.2.1 Questionnaire

The survey questionnaire was developed to investigate Korean English teachers' self-efficacy of LAL and their self-identification in the teaching profession. In order to assure the validity of the survey, an expert in English assessment and Applied Linguistics offered guidance during the construction of the questionnaire and it was piloted with three in-service

Korean English teachers. These teachers provided feedback on the wording, the structure, and alternative interpretations of the questions. The survey questions mainly used a five-point Likert scale and there were open-ended questions as well. Forty-nine items were surveyed in total. The questionnaire is provided in Appendix 1.

The questionnaire was structured in two parts. Part A consisted of questions meant to measure teachers' self-efficacy of LAL, while Part B included questions intended to reveal NNEST identity. To begin with, Part A was an altered version of the teacher's assessment professional competency test by Song and Kim (2007). Other surveys for secondary English teachers' assessment competency were also referred to in the construction of survey questions (Jeon & Oh, 2006; Vogt & Tsagari, 2014). Since the questionnaire of Song and Kim(2007) was aimed to investigate teachers' LAL along the entire process of assessment, it covered five domains: 'Choosing methods for assessment,' 'Developing instrument for assessment,' 'Grading and scoring,' 'Using assessment results' and 'Knowing ethics of assessments.' The present study, however, aims to investigate difference in perceptions of teachers' LAL competence in assessing different language skills such as listening, reading, speaking, and writing. Therefore, the questionnaire has been reorganized into four sections as follows: assessing listening skills, assessing reading skills, assessing speaking skills, and assessing writing skills. The five

domains of LAL are investigated in each language skill separately. For instance, in the first section, questions asked about the five domains in assessing listening skills along the process of assessment from planning to assessment ethics. Part A of the survey is composed of four sections each of which investigates teachers' self-efficacy in each language skill. Each section has eight questions and the questions were evenly extracted from the five domains of assessments (Song & Kim, 2007). Each question used a five-point Likert scale, which ranges from 'Strongly Disagree' to 'Strongly agree'. Response categories were coded to numbers before the data analysis; 1 for 'Strongly Disagree', 2 for 'Disagree', 3 for 'Neutral', 4 for 'Agree', and 5 for 'Strongly agree'. Therefore, mean scores indicated participants' self-efficacy of assessments. Higher mean scores implied positive self-efficacy, while lower mean scores pointed to negative self-efficacy of teachers in assessment practices. Each section has an open-ended question asking about difficulties in their assessment practice. Responses from the open-ended questions were collected for thematic analysis to reveal the difficulties teachers encounter. Part A consisted of thirty-six items in total, including four open-ended questions.

Part B, with 13 items, focused on teacher identity. Questions about NNEST identity were composed based on the previous studies of NNEST identity by Aneja (2016), Choe (2008), Hong (2013), and Hwang and Lee

(2018). Twelve items were using a five-point Likert scale and one open-ended question. Response categories were coded the same as part A, but five items (item 44 to item 48) were reverse coded. Those items were reverse coded in order to ensure the accurate measurement of KETs' identity and also to keep respondents from answering carelessly. In this part, a higher mean score indicated teachers' positive NNEST identity and a lower mean score showed a negative NNEST identity as a less qualified teacher. The open-ended question asked about teaching and assessment difficulties caused by the teachers' own language proficiency. The responses were analyzed thematically to examine the factors that influence NNEST identity. Survey questions are provided in Korean to eliminate any possible misinterpretations.

3.2.2. Semi-structured Interview

To elicit more detailed reflections on teachers' self-efficacy in their assessment experiences as well as their professional identity in language teaching classrooms, a qualitative investigation was essential. This study employed a semi-structured interview as an additional method in order to derive a richer dataset, demonstrating the interrelationship between teachers' assessment practices and identity construction.

A total of ten teachers who voluntarily agreed to participate in the

interview were randomly selected as interviewees. All the interviews were conducted in Korean and were guided by five predetermined questions to encourage the participants to elicit and elaborate on their perceptions of language assessment and self-identification as a teacher and a assessor. The questions were developed with reference to the teacher identity profiles from Aneja (2016), Choe (2008), Hong (2013), and Hwang and Lee (2018) and they were revised multiple times regarding the pilot study responses and feedbacks from two in-service teachers who did not participate in the interview. The interview questions are provided in Appendix 2.

3.3 Procedures

This section presents the procedures of data collection and analysis. Section 3.3.1 provides information about the data collection process. In Section 3.3.2, methods employed for data analysis are provided.

3.3.1 Data Collection

The data collection started after gaining approval from the Ethics Committee of the Institutional Review Board (IRB) of Seoul National University. Seventy in-service Korean English teachers who teach in middle

or high schools were recruited as participants from an online community of in-service Korean English teachers. An online link to LAL Self-efficacy and Teacher identity survey was sent to participants who voluntarily agreed to participate in the research. A description of this research was provided to participants online and the queries raised by the participants were answered by the researcher. The researcher also notified participants about their right to discontinue participating in the research at any time. The questionnaire was delivered online via Google Survey form and data was collected for ten days.

After analyzing survey response data, nine participants were chosen for a follow-up interview. The interview was conducted individually through an online video conference or phone call by the choice of the interviewee. The researcher provided a description of the study and notified participants about the recording of the interview, and the right to withdraw their consent of participation at any time. Every participant agreed to be interviewed voluntarily and completed the consent form. The interview took about twenty minutes, and the interviews were audio-recorded and later transcribed.

3.3.2 Data Analysis

Both quantitative and qualitative analyses were used to address the research questions. Data from the survey was analyzed statistically by SPSS

27. Interviews were transcribed and analyzed by thematic analysis.

The data from the Teacher LAL self-efficacy and teacher identity survey questionnaire were analyzed with SPSS 27. A descriptive analysis of the survey responses from Part A was calculated to measure teachers' self-efficacy of language assessment literacy. As the survey used a 5-point Likert scale, the responses were categorized from 1 ("Strongly Disagree") to 5 (Strongly agree). The responses from Part B are also analyzed descriptively. Five items (item 44 to item 48) were reverse coded before the analysis. The means and standard deviations were measured on each item, and in each section. Then, the questionnaire's internal consistency reliability was measured by the calculation of Cronbach's alpha coefficient. A repeated measures ANOVA was carried out to compare the participants' LAL self-efficacy in four skills of language: Listening, Reading, Speaking, and Writing. Then, to measure the correlation between teachers' LAL self-efficacy and their teacher identity, a Pearson correlation coefficient between Part A and Part B was calculated. The results of the statistical measure helped identify the difference in LAL efficacy between domains of language skills in assessing language, and the relation between LAL efficacy and teacher identity.

The response data from open-ended questions and the transcribed interview data were analyzed qualitatively in order to uncover sociocultural

and contextual factors that influence the LAL efficacy of teachers and also the correlation between teachers' LAL self-efficacy and their identity formation. The thematic analysis was employed on the basis of the inductive coding process (Creswell, 2014). Responses from open-ended questions were thematically analyzed and categorized into keywords. Recordings of interviews were transcribed as a first step and guided by thematic analysis (Braun & Clarke, 2006; Creswell, 2014), the data was analyzed. The data were read and initial codes were generated by open coding to discover recurring themes and subthemes. The same data were examined repeatedly and codes were added, merged, and removed to redefine the themes. After the final themes and subthemes were identified, specifics of each theme and subthemes were elaborated. Then, the themes were organized with the reference to the framework for teacher identity investigation by Trent (2015). According to the framework, the themes were categorized into two constituents of identity: "identity-in-discourse" and "identity-in-practice" (Trent, 2015, p.46). Then, the themes and the subthemes were analyzed in detail. In the following chapter, the findings from the results of the analysis will be presented.

CHAPTER 4.

RESULTS

This chapter presents the findings of the study. The former two sections describe the results of the survey with quantitative analysis while the latter two sections provide the qualitative results from the data collected from the open-ended questions of the survey and the interviews. Section 4.1 reports on the findings regarding the KETs' self-efficacy of LAL in different language skills. Section 4.2 reveals the correlation between KETs' self-efficacy and teacher identity. Section 4.3 presents recurring themes of contextual factors affecting KETs' assessment practice and teacher identity formation. Section 4.4 delineates an in-depth description of the relation between teachers' assessment efficacy and the teacher identity.

4.1 Teachers' Self-efficacy of LAL

With regard to the first research question, KETs' self-efficacy of LAL was examined through the Part A of the survey questionnaire. Statistical analyses were performed using SPSS 27 for Windows. Part A consisted of

four sections which measures KETs' self-efficacy of LAL in each language skill. Each section has eight items that use a 5-point Likert scale. Cronbach's alpha for each section was from .814 to .911, indicating a high level of reliability. The reliability and descriptive statistics for the four sections are presented in Table 4.1.

Table 4.1

KETs' LAL Self-efficacy in Four Skills (N=70)

Section	Cronbach's α	Mean	SD
Listening	.855	3.82	0.539
Reading	.911	3.90	0.639
Speaking	.901	3.85	0.629
Writing	.903	3.95	0.584

As can be seen in Table 4.1, the mean score of KETs' self-efficacy of LAL differs depending on the language skills. Since each section has eight items with a five-point Likert scale, the maximum total score of each section is forty. Participants felt most confident about their LAL in the area of assessing writing (M=3.95, SD=0.584) and felt least confident when they assess listening skills (M=3.82, SD=0.539). The mean score for assessing reading skills (M=3.90, SD=0.639) is relatively higher than assessing speaking skills (M=3.85, SD=0.629). Table 4.1 shows that KETs feel more confident in assessing reading and writing than assessing listening and

speaking. Figure 4.1 illustrates the difference in scores in each section of LAL self-efficacy in the boxplot.
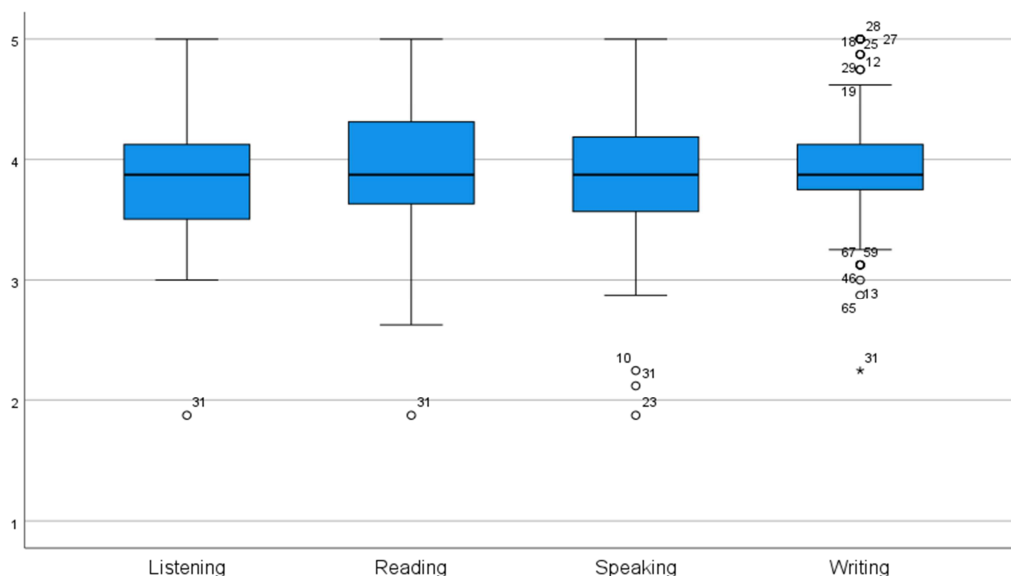
Figure 4.1 Boxplots of KETs' LAL self-efficacy in four skills

Figure 4.1 indicates that in assessing speaking, participants show a slightly higher level of confidence than assessing other skills. It does not correspond to the difference in mean scores in Table 4.1. Hence, the descriptive analysis of the results implies that the minute differences in the four sections do not seem to be statistically significant. Also, the boxplots illustrate that a majority of KETs expressed a consistent tendency of high self-efficacy in assessing writing. On the contrary, it seems that in assessing reading and speaking skills, the confidence varies between KETs. The result

slightly disagrees with the previous research results that reported EFL teachers expressed a strong need for LAL training in assessing productive skills (speaking and writing skills) even though they had received training most in the area (Chung & Nam, 2018; Vogt & Tsagary, 2014).

To examine whether these differences in mean scores were statistically significant, a repeated measures ANOVA was conducted. The spherical assumption was checked by the Mauchly spherical test and the result shows a p-value of .177 (p>.05), so the assumption of sphericity is observed. Table 4.2 shows the result of the repeated measures ANOVA.

Table 4.2.

Repeated measures ANOVA results

Source of Variation	SS	df	MS	F	Sig.	Partial η^2
Within skills	44.268	3	14.576	2.047	.108	.029
Error	1491.982	207	7.208			
Total	1536.250	210				

The null hypothesis is that there is no difference in participants' LAL self-efficacy scores in sections of assessing different language skills was evaluated. The repeated measures ANOVA result determined that there is no significant difference in the mean scores across sections of four skills

(F(3,207)=2.047, p = .108). The result of the within-subjects effect shows a *p*-value greater than 0.05. It indicates that it fails to reject the null hypothesis (H_0: $\mu_1 = \mu_2 = \mu_3 = \mu_4$) and the mean score difference between sections cannot be considered statistically significant. Therefore, the statistical analysis of Part A suggests that the differences in mean scores in sections cannot be considered statistically significant and it indicates that there was no significant difference across assessing the four different language skills in participants' LAL self-efficacy.

4.2 LAL Self-efficacy and Teacher Identity

This section presents the results of survey Part B about KETs' NNEST identity and measures the correlation between KETs' LAL self-efficacy and their identity by conducting a Pearson correlation analysis.

Table 4.3 presents the descriptive statistics and Cronbach's alpha for KETs' NNEST identity. The Part B of the survey consisted of twelve items with a 5-point Likert scale, so the maximum score is sixty. The high score means a more positive NNEST identity construction of the respondent, while the low score indicates a negative NNEST identity. The mean value for KETs' NNEST identity was 3.37 at a standard deviation of 0.620 and the Cronbach alpha is .814, which shows a high level of reliability.

Table 4.3

KETs' NNEST Identity (N=70)

Section	Cronbach's α	Mean	SD
Identity	.814	3.37	0.620

In order to examine whether there is a statistically significant relationship between the KETs' LAL self-efficacy and NNEST identity, Scatter-plots were drawn to check the linear relationship between variables. The Scatter-plots in Figure 4.2 support a positive correlation between the variables. The scores of LAL self-efficacy from each section are designated on the axis (Y) and the NNEST identity on the axis (X). This demonstrates that the patterns of overall scores of each section are close to a straight line, and a linear association between KETs' LAL self-efficacy and their NNEST as a professional teacher identity.
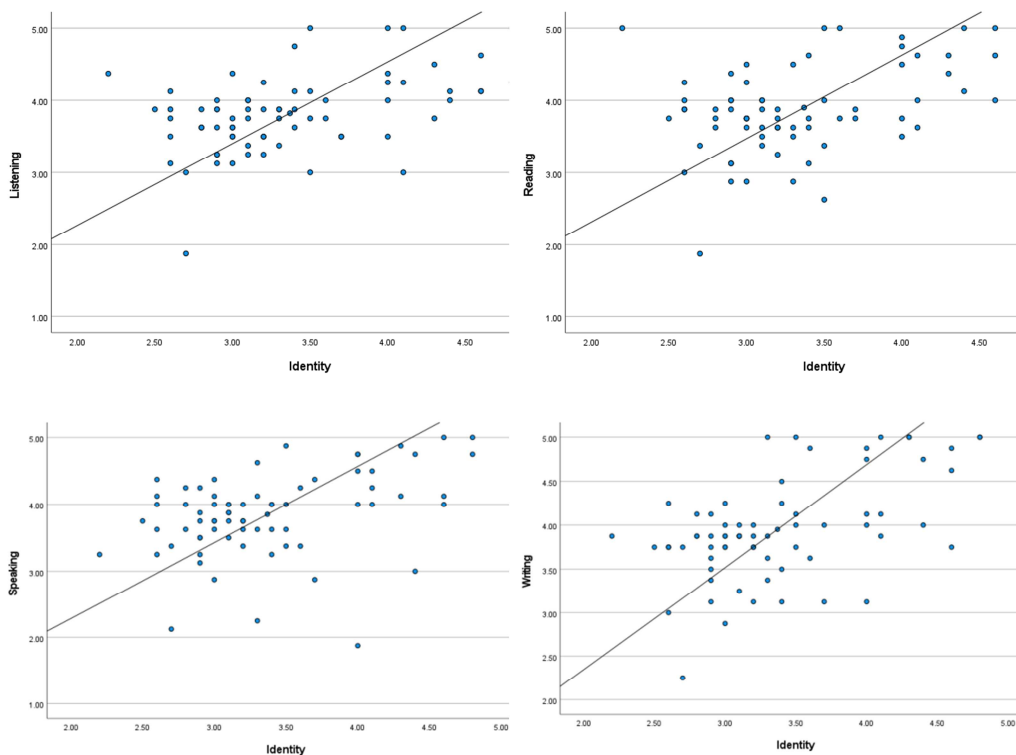
Figure 4.2 Scatter-plots of LAL self-efficacy in each language skill and NNEST identity

Then, the Pearson coefficient was calculated. The results are provided in Table 4.4.

Table 4.4

Results of the Correlation (N=70)

		Listening	Reading	Speaking	Writing
NNEST identity	Pearson Correlation	.468**	.480**	.450**	.570**
	Sig.(2-tailed)	.000	.000	.000	.000

**. Correlation is significant at the 0.01 level (2-tailed).

The results in Table 4.4 indicate that a statistically significant positive correlation exists between each part of LAL self-efficacy and NNEST identity at the level of .01. It can be interpreted that if the teacher participant is with a positive NNEST identity, he/she is expected to express a high level of LAL self-efficacy. The Pearson correlation coefficient of the LAL self-efficacy score for the assessing writing section and the identity ($r = .570$, $p<.01$) was the highest among the variables of the four sections; followed by the score of the reading section ($r = .480$, $p<.01$), the score of the listening section ($r = .468$, $p<.01$), and the speaking section was the lowest ($r = .450$, $p<.01$). According to Plonsky and Oswald (2014), in the field of language education research, r values under .25 can be considered small, .40 medium, and close to .60 large. Hence, the strength of correlation between listening, reading, and speaking sections and NNEST identity could be interpreted as a medium, while the correlation between writing and the identity can be interpreted as large.

The Pearson coefficient statistically revealed that the higher the KETs' LAL self-efficacy is, the more positive NNEST identity they possess, and the lower their LAL self-efficacy is, the less positive NNEST identity they expressed.

4.3 KETs' Assessment Practice and Difficulties

The thematic analysis was conducted on the open-ended question responses from the survey to find what KETs perceive as effective factors in their assessment practice and also difficulties KETs encounter related to their language proficiency. The responses from open-ended questions were repeatedly read and grouped into categories of themes which were inductively constructed. The themes were continuously reexamined and reorganized during the analysis. Subthemes were extracted from each theme to reveal major difficulties teachers experience when they assess at schools. Lastly, for better understanding, the themes were grouped into three categories: Developing instruments for assessment, conducting an assessment, and using assessment results.

Some themes re-occurred continuously in all four sections: assessing listening, reading, speaking, and writing. First, KETs were mainly concerned about maintaining the validity and the reliability of the assessment. Most of the participants agreed that these issues appeared as major and most problematic aspects of school assessment regardless of which skill is to be tested. "Constructing valid and reliable scoring criteria" and "making effective and valid test items" were most mentioned as the areas that teachers felt most difficult to fulfill. Excerpt 4.1 to 4.6 below reveals KETs' such concerns.

Excerpt 4.1. Because of the performance-like nature of speaking, lots of variables influence learners' speaking behavior. So it is hard to establish clear and precise scoring criteria. Because of that, scoring rubrics for speaking assessments tend to become too simplified and it encourages students to memorize prewritten scripts when they prepare for speaking tests.

Excerpt 4.2. It was hard to maintain the inter-rater reliability in speaking assessments. Even though teachers established a scoring rubric before the assessment, there were differences between teachers in the actual grading.

Excerpt 4.3. There were lots of conflicts between teachers in the process of developing scoring criteria for listening assessments. For example, giving minus points for minor mistakes like spelling errors was highly debated.

Excerpt 4.4. It is hard to measure the actual reading skills of students in school reading assessments because only the textbook materials are taught and tested.

Excerpt 4.5. It seems doubtful whether school reading assessments are valid for measuring students' reading ability because the exam questions are written only from textbook readings for preventing the prevalence of private English lessons in advance.

Excerpt 4.6. School tests mostly consist of multiple-choice questions for the sake of reliability and practicality of assessments. But I'm not sure whether the multiple-choice questions can be a valid method for reading assessments.

Moreover, the responses suggest that KETs were aware of the interrelation of these features. They are concerned that when the construction of a rubric is inclined to secure the reliability of the test, essential domains of language ability could be overlooked in the assessment and it might result in a deterioration of the validity of the assessment. Such concerns were found in all four sections. The following excerpts illustrate these concerns.

Excerpt 4.7. In grading writing, scoring criteria that we use in schools are generally concerned with grammar and vocabulary to avoid the raters' subjective scoring judgments. Therefore, the overall

organization and contents of the writing are overlooked in school writing assessments.

Excerpt 4.8. To raise the inter-rater reliability, scoring criteria for speaking assessments are developed to reject a subjectivity of a scorer. Therefore essential constituents of speaking skills like fluency or confident attitude are less likely to be included in rubrics.

As the excerpt 4.7 and 4.8 show, KETs mentioned that because of the efforts to make tests objective and impartial, current assessments and scoring criteria undermine ambiguous constructs such as attitudes of a speech or an organization of writing. They considered current assessments are focused on more objective aspects such as grammatical errors or language misuse.

In addition, KETs were worried about integrating classroom lessons and assessments. KETs expressed concerns that some language skills are not fully taught in class and yet are subject to be assessed and excerpt 4.9 and 4.10 reveals such concerns.

Excerpt 4.9. I feel uncomfortable when I do speaking assessments because I think speaking lessons are not provided enough in school classes.

Excerpt 4.10. Because we use the listening tests made by external institutions, the tests are not related to what is taught in class. And the school English lessons do not include many listening lessons; it seemed that there is a disparity between the assessments and learning in school English classes.

Also, the practicality of the test appeared to be problematic due to a large number of students. An extreme level difference between students was also mentioned as a perplexing problem in assessing four skills. Lastly, KETs are concerned about insufficient feedback given to students in school assessments due to a large number of students and in-class time limitations. Excerpt 4.11 and 4.12 illustrate those problems KETs experienced.

Excerpt 4.11. In writing assessments, I try hard to provide constructive feedback to every student but it takes too much time and effort because of the large number of students. I teach about 170

students. Sometimes it is hard to just score their drafts.

Excerpt 4.12. The difference in students' speaking ability is extreme between advanced and lower-level students. Some students can talk fluently in English while other students cannot even say a sentence. So it was difficult to assess them with an appropriate assessment method. Also, speaking tests were too time-consuming as it has to be done one by one.

In assessing the listening skills of learners, seven themes were revealed as sources of difficulties as seen in Table 4.5. In assessing listening, securing validity was a major issue. According to the response, listening tests made by external organizations of English teachers are widely used for assessing listening skills in schools as teachers rarely know how to develop a listening test other than traditional multiple-choice questions or dictations.

Table 4.5

Identified Themes and Subthemes in Assessing Listening Skills

Assessing Listening		
Sections	Themes	Subtheme
Developing instrument for assessment	1. Validity	a. Lack of content validity in test items b. Using tests made by an external agency
	2. Teachers' limited knowledge in test development	a. Difficulty of self-production of recording by teachers b. Unawareness of listening test formats other than dictation and multiple-choice questions
	3. Broad proficiency range of students	Difficulty of developing assessment items at an appropriate level for students
Conducting an assessment	4. Practicality	a. Securing an environment conducive to taking the test b. Retaining equal setting for every test takers
	5. Setting evaluation criteria	Conflict in setting evaluation criteria in dictation tests
Using assessment results	6. Relation between class and assessment	Lack of listening lessons in class
	7. Providing feedback	a. Difficulty of using test results as diagnosis b. Insufficient individual feedback for students

Moreover, they are concerned if they produce their own recordings for a listening test, complaints about intelligibility and pronunciation of the listening material will be raised due to its non-nativelikeness. KETs expressed unease with using the premade tests as well, as the premade tests are not related to the in-class listening lessons and also cannot reflect learners' current level of listening skills. Excerpt 4.13 reveals these concerns.

Excerpt 4.13. When we use premade listening tests it is hard to find a valid test that is appropriate for the students' listening abilities. But I cannot even think of making level-appropriate listening assessments by myself because it is going to be too difficult. For example, if teachers make their own recordings, many test-takers would complain that "The pronunciation and accents of the listening test were not like those of native English speakers."

Some participants reported that developing their own listening tests with textbook material enabled more meaningful assessment.

In addition, the practicality of the test was also noted as a troublesome area. In order to prevent the leaking of test questions, listening tests need to be conducted simultaneously for students in the same grade.

Otherwise, tests with the same level of difficulty should be developed. Excerpt 4.14 illustrates these in detail.

Excerpt 4.14. When a listening assessment is not conducted at the same time for the whole grade, the test items can be revealed in advance to the test takers who take the test later. In this case, the assessment would lose its fairness and integrity. If the assessment cannot be conducted at the same time, teachers have to prepare several sets of tests with an identical level of difficulty, which is not an easy task. I felt this issue was the most difficult one to deal with in assessing listening skills.

Table 4.8 presents six themes found revealed as sources of difficulties in assessing reading skills of learners. The open-end question responses in the assessing reading session indicates that KETs expressed more confidence in assessing reading than assessing any other skills. A few participants directly mentioned that "It is the most confident area of language assessment.", or "Relatively, reading assessment seems to be the most plausible and reliable language assessment in schools."

Table 4.6

Identified Themes and Subthemes in Assessing Reading Skills

	Assessing Reading	
Sections	Themes	Subtheme
Developing instrument for assessment	1. Validity	a. Lack of construct validity and face validity b. Test materials limited to textbooks c. Conflicts between retaining the validity of the test and the need to differentiate the English levels of student
	2. Lack of resource	Difficulty of finding appropriate reading material for assessment
	3. Teachers' limited knowledge in test development	a. Difficulty of eliminating possible ambiguity of interpretation in making test items b. Limited format of test
	4. Broad proficiency range of students	Difficulty of developing assessment items at an appropriate level for students
Conducting an assessment	5. Complex nature of reading comprehension	Difficulties in measuring reading ability due to its broad spectrum and complexity
Using assessment results	6. Providing feedback	Insufficient individual feedback for students

Some participants, however, raised the issue of validity. KETs revealed concerns in terms of construct validity as most reading tests cannot measure the reading strategy use of students and are mainly constructed with reading comprehension questions. Also, test formats that are limited to multiple-choice questions raised an issue of validity as choosing right answer does not guarantee the measurement of students' reading ability. More importantly, since exam questions are limited to textbook reading materials which are already taught in class, KETs raised doubt about whether those reading tests are capable of measuring the actual reading ability of students. Excerpt 4.15 presents this doubt of a KET.

Excerpt 4.15. Generally, reading assessments are conducted in the form of paper-based written tests. And mostly, they are focused on the knowledge of the textbook materials and grammatical knowledge. However, the gap of language ability between students I teach is huge and many of them are deficient in English. It makes me doubt that this format of reading assessments can be meaningful or valid for my students. So, I've been trying to use an alternative form of reading assessments such as filling in blanks of graphic organizers.

In addition, the responses show that KETs are aware that reading

requires students' complex knowledge of vocabulary, context, and syntactic understanding of the text and it is hard to develop appropriate test items that measure reading ability in a comprehensive approach.

In assessing speaking skills, nine themes were found as shown in Table 4.7. In assessing speaking, teachers most frequently mentioned about difficulties of making speaking tests and scoring criteria valid and also reliable. Since speaking behavior is a performance that should be scored instantly, many teachers express the difficulty of scoring speaking assessments, as well as the strong need for clear and objective scoring criteria. Yet, they were concerned about the validity of the speaking test as well. Some KETs said that for the sake of reliability and objective criteria, essential domains of speaking ability, such as fluency, are excluded or minimally included in the speaking test rubrics.

Table 4.7

Identified Themes and Subthemes in Assessing Speaking Skills

Assessing Speaking		
Sections	Themes	Subtheme
Developing instrument for assessment	1. Validity	a. Lack of content validity in scoring criteria because of raising the reliability of the test b. Degeneration of the speaking test into testing memory c. Difficulty of planning valid speaking task
	2. Reliability	a. Raising inter-rater reliability b. Difficulty of scoring due to performance-like nature of speaking behavior
	3. Broad proficiency range of students	Developing adequate speaking tasks and criteria that can include every student with different proficiency
Conducting an assessment	4. Practicality	Demanding too much time
	5. Instant nature of speaking activity	a. Difficulty of scoring the performance immediately b. Difficulty of securing consistent and precise grading
	6. Affective filter	Dealing with students' public speaking anxiety
	7. Securing fairness of the test	Complaint report relating to grades due to ambiguous interpretations of spoken forms
Using assessment results	8. Providing feedback	Insufficient individual feedback for students
	9. Relation between class and assessment	Lack of in-class speaking lessons

Moreover, it has been pointed out that speaking assessments are sometimes perceived as measurements of memory rather than measuring speaking ability as students generally prepare for their tests by memorizing the whole script in advance. This tendency seems to be widespread in schools, according to the responses. Many KETs are concerned about this degeneration of speaking assessments as excerpt 4.16 and 4.17.

Excerpt 4.16. Because of the large number of students, most speaking assessments are conducted in the format of a short individual presentation. So students are supposed to prepare a script for the presentation in advance. I was doubtful whether this type of speaking assessment can be valid for measuring students' actual speaking skills.

Excerpt 4.17. In the school I teach, the majority of the students are not good at speaking English. So eventually, speaking assessments becomes a memorization test for them.

Furthermore, disputes about the fairness of test results were also concerned. A response mentioned that "Because of distrust of non-native English teachers' scoring ability on speaking assessments, complaints on

speaking test results are frequently reported on schools." A few responses directly mentioned that assessing speaking is the most difficult and the least confident area of assessment. Overall, the responses implied that KETs are struggling most in assessing speaking skills.

Eight themes were revealed in writing assessment section as presented in Table 4.8. In assessing writing, some KETs mentioned that writing assessments are also degenerated as memorization tests, like speaking assessments, especially for students with lower proficiency. Other KETs noted that using process-based writing assessments was more beneficial for students as it keeps students from memorizing their pre-written drafts. Still, they were worried about assessing students with lower proficiency levels as those lower-level students cannot write a paragraph.

Table 4.8

Identified Themes and Subthemes in Assessing Writing Skills

Assessing Writing		
Sections	Themes	Subtheme
Developing instrument for assessment	1. Validity	a. Lack of validity in scoring criteria to raise reliability of the test b. Degeneration of the writing test into testing memory
	2. Reliability	a. Raising inter-rater reliability b. Difficulty of developing objective and clear scoring criteria that fits achievement criteria.
	3. Involving additional writing tools	Incorporating the use of online writing tools
Conducting an assessment	4. Practicality	Difficulty in scoring due to the large number of students
	5. Broad proficiency range of students	Difficulty of involving students with lower proficiency
	6. Difficulty in grading	Difficulty in grading with consistency due to various possible writings for a single task
Using assessment results	7. Providing feedback	Insufficient individual feedback for students
	8. Relation between class and assessment	Lack of in-class writing lessons

Developing valid and reliable scoring criteria was also revealed as a major issue. According to responses, a tendency to use analytic rubrics was prevailing as a countermeasure for securing inter-rater reliability, but many KETs expressed concern about overlooking significant aspects such as assessing content, cohesion, and coherence of writing. Following excerpt 4.18 illustrates such concern.

Excerpt 4.18. I think teachers need scoring criteria which correspond to the objective of assessments. For example, when the objective of a writing assessment is measuring students' communicative use of language in a written form, mechanical errors or grammatical domains of grading should be regarded as less important. However, in school writing assessments, generally those two aspects are more focused.

Also, KETs implied that they attempt to use various forms of testing tools in assessing writing. For example, the responses included KETs' experiences of using online translator, process-based writing, or holistic rubrics for writing assessments and relating difficulties they had.

Lastly, in the last section, KETs were asked about difficulties they

faced in their teaching and assessing due to their language proficiency, and

six themes were found from the responses. Table 4.9 summarizes the themes.

Table 4.9

Identified Themes and Subthemes about KETs' Teacher Competence

Difficulties KETs had due to language proficiency	
Themes	Subtheme
1. Teachers' speaking ability	a. Lack of certainty in deciding correctness of an expression when they speak b. Difficulty of using various and detailed English expressions in their speech c. Lack of confidence in intonation, stress, and pronunciation d. Difficulty of providing immediate answer to students' questions about English usage
2. Checking accuracy	Lack of native-like intuition in checking grammaticality in written forms
3. Authenticity	a. Difficulty of using appropriate collocation and idioms b. Reflecting dynamic change of English usage in the real world to in-class lessons.
4. Comparison of language proficiency between English teachers	Constant comparison of teachers' English proficiency in their teaching and making assessments
5. Teaching ability other than linguistic proficiency	Lack of experience in various forms of language assessments
6. Insufficient teacher training	Reduction of qualified trainings for English teachers

According to the responses, it seems that KETs were concerned about their speaking ability when they teach and assess at schools. Mostly, KETs felt the need to work on their speaking ability when students suddenly ask them to provide adequate English expressions as an example. They mentioned that at those moments, they cannot be certain whether the expression they come up with is appropriate or not. KETs are likely to attribute this uncertainty to a deficiency in their linguistic ability. When a KET co-teaches with a NEST, this inclination seemed to increase. Excerpt 4.19 below reveals such a tendency.

Excerpt 4.19. In a co-teaching class with a native English-speaking teacher, I gave an answer to a student's question and the native teacher immediately corrected my answer. Since I'm not a native speaker, it is difficult to convey the exact nuances in expressions. Generally, the difference in linguistic proficiency was not my concern, but it is the part that I feel sorry for my students.

Some participants shared that they felt their speaking ability is insufficient when they taught a student who had lived abroad. Also, other responses showed KETs' lack of confidence in intonation, stress, and

pronunciation. KETs also noted that they are studying by themselves to improve their speaking proficiency so they can deal with those circumstances. Other KETs, on the other hand, said that they think their speaking proficiency is good enough for classroom talks. They argued that other than linguistic proficiency, teaching ability such as using the diverse and appropriate types of assessments or explaining linguistic aspects efficiently to students is much more important for teachers.

Moreover, the responses revealed that when it comes to writing assessments or grading open-ended questions in tests, KETs faced difficulty in determining acceptable answers. According to some responses, KETs spend a lot of time searching online to figure out whether a certain expression is used in authentic circumstances in order to decide its correctness. Also, some teachers mentioned experiences that sometimes students who had lived in inner circle countries file a complaint about their grades as they insist the ungrammatical expressions they use are actually natural. And they added these moments are difficult to deal with and they felt insufficient of their linguistic proficiency at those times. Overall, the responses imply that teachers' linguistic proficiency is considered an influential factor in their teaching and assessment to varying degrees.

4.4 KETs' Identity in Teaching and Assessing Language

This section presents the findings of the thematic analysis of transcribed interviews. Nine KETs enrolled in the in-depth interview after their survey participation. The analysis was conducted to examine influential themes regarding KETs' professional identity construction and the influence of the identity work in practicing language assessment. The analysis of interviews was based on the framework of the teacher identity investigation by Trent (2015).

The framework suggests that identity construction reflects both discursive and experiential influences in multi-dimensions of intrapersonal, interpersonal, and institutional levels. Those two constituents are categorized as identity-in-discourse and identity-in-practice. In the discursive construction of identity, individuals make commitments in terms of modality and evaluation. Modality refers to individuals' commitment to what they perceive as truth and obligation, and evaluation involves what individuals believe to be desirable or undesirable. In the experiential construction of identity, three modes of belonging were suggested: engagement, imagination, and alignment. Engagement describes that individuals engage in certain communities and learn the conventions in those communities. In the mode of imagination, they gain membership of imagined communities and legitimize their identity positionality. Alignment indicates that individuals coordinate

their activities within a broader context of organizations. Based on these two categories of discourse and practice, the themes from the interviews were analyzed and organized. Figure 4.3 summarizes the themes from micro to macro dimensions.

Discourse	Agency	Practice
Intrapersonal: Contrasting identity of teacher expert and learner		**Engagement:** Teacher collegiality
Interpersonal: Constant tension against stigmatization		**Imagination:** Imagined membership of EFL teachers
Institutional: Priority on test objectivity		**Alignment:** Aligning with school circumstances and fulfilling the needs of the students

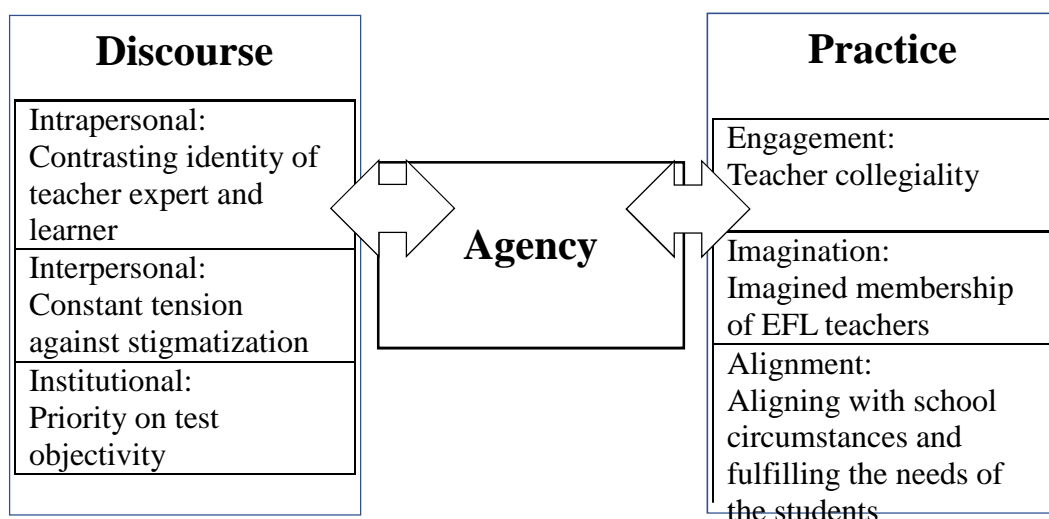**Figure 4.3. Identified themes in KETs identity construction**

4.4.1. Teacher Identity in Discourse

Through language and discourse, individuals construct, maintain, and negotiate their identities to a meaningful extent (Varghese et al., 2005). At the intrapersonal level of discourse, they locate themselves in a discourse explicitly through the overt naming of their identity. According to interviews,

several KETs identified themselves as "language education experts," which was legitimized through reference to their teaching experiences and qualifications. 4 KETs (Teacher A, B, C, F) specified by claiming that facilitating students' learning in a class needs complex integration of various teaching skills and methods, hence, the integrated ability gained through continuous trial and error in hands-on teaching experience justifies teachers to be recognized as experts. The act of identification is also explicitly revealed when some KETs differentiate their self-claimed identity position from NESTs in schools, as they mention the NESTs' lack of in-class teaching ability.

KETs, on the other hand, also identified themselves as "language learners." Many KETs pointed out that their prior experience as language learners influenced their teaching and assessment practice in both positive and negative ways. For instance, teachers mentioned that as they learned English themselves, they can understand specific difficulties students would face, and can easily lead them to avoid potential pitfalls and support successful language learning. Some teachers, on the other hand, noted that they were immersed in the test formats they experienced as students, such as multiple-choice questions and semi open-ended questions, and this made it difficult for them to try out new assessment methods. In addition, teachers stated that they are constantly working on their own language proficiency or

they feel the strong need to do so in order to reach the expected language proficiency, and it indicates that they identify themselves as perpetual language learners.

KETs' learner identity was obviously demonstrated in assessing speaking and writing skills. Most interviewees stated that speaking was the most difficult skill to evaluate due to the instant nature of speaking performance, and some KETs shared that they feel a burden in making decisions to grade students' fluency and accuracy immediately after their speaking.

They also mentioned that in assessing speaking and writing, they feel hesitation when they give a mark to the acceptability of certain language use and some teachers explicitly mentioned their non-nativeness as a source of such uncertainty. In excerpt 4.20, Teacher A illustrates a lack of self-confidence in assessing productive skills:

Excerpt 4.20. Even as a language learner, I did not pay much attention to developing speaking and writing skills, so it seems that the time and effort I invested in productive skills was relatively small. I think that in learning English, the areas that I haven't mastered yet are endless no matter how much effort I put into them. And also,

languages are constantly changing and evolving so I often have doubt on a language use whether it is possible or not. Even if I look it up every time, I sometimes cannot easily get a definite answer, so I feel less confident in assessing speaking and writing. (Teacher A)

In the excerpt, teacher A's self-awareness as a perpetual learner of English is unveiled. On the other hand, Teacher A also identifies herself as an expert English teacher in excerpt 4.21 below.

Excerpt 4.21. I consider myself an expert in English education. It is not based on the fact that I have so-called 'native-like intuition' more than others. I think of myself as an expert because I realize that in order to provide meaningful language lessons to students in class and to facilitate their learning, very sophisticated techniques are needed with integration of educational knowledge and teaching experiences. As I manage to accomplish such tasks and teach students, I feel that I'm an expert. (Teacher A)

Hence, KETs seem to have contradicting identities both as an expert and a learner. This contradiction is also revealed in other teachers' responses

(Teacher B, C, D, and Teacher H) in different degrees. The intrapersonal conflict of KET's identity is shared by Teacher B in excerpt 4.22:

> *Excerpt 4.22. I think working on English proficiency is very important for English teachers and if you have native-like English fluency, it could be a unique weapon, and a big advantage as a teacher. I used to think that the language ability is the most important qualification for a teacher, but on the other hand, as I teach in middle schools I realized that technical aspects of teaching such as delivering the lesson contents to students in not only interesting but also accurate way and teaching a class with adequate control and lead are also very important. Still, I think English ability is a part that I have to work on continuously. You can never learn too much. But unless you are a native speaker, in fact, there is a limit to the ability you can have, so without losing courage, I would like to focus on developing a teaching method. (Teacher B)*

As illustrated in the excerpt above, mixed identity as a language learner and also a teaching expert coincides as KETs recognize themselves as NNESTs. Due to this coexistence of mixed and contradictory self-concepts, they show anxiety and self-doubt about their own language during teaching

and assessment. In the open-ended responses and also in the interviews, the anxiety that KETs cannot be certain about adequate and acceptable use of English was revealed as one of the major difficulties teachers face in assessment practice. Some respondents pointed out "lack of native speaker's intuition" as a specific deficiency they had when they assess students' language behavior, especially for checking linguistic accuracy.

This anxiety is found not only on an intrapersonal level but also on an interpersonal level. Some KETs reported that they are afraid of making linguistic mistakes in class because they might be seen as their language proficiency is incompetent by students. Some of them even specifically pointed out that when they teach students from abroad, they feel anxious about their pronunciation. This kind of insecurity exists not only between teachers and students but also among teachers. Teacher C noted the "defensive attitude" of KETs and attributed the cause to the KETs' intrinsic uncertainty as below:

> *Excerpt 4.23. I often felt the self-defensive attitude from co-working teachers in the process of assessment and it is one of the hardest parts I experienced as a teacher. I and my colleague teachers talked about it and we agreed that it comes from teachers' internal uncertainty on*

English use. Since English is a foreign language to all KETs, teachers

cannot be completely sure of what they already know. Because of this

uncertainty, when others raise any other opinion on what teachers

said or taught they generally react in a defensive manner rather than

mediating between interpretations. KETs seem to presuppose that they

should know well about linguistic matters or else they would be

considered disqualified. So when KETs face disagreements, they tend

to accept it as if someone told them that they are wrong, and they are

likely to become self-defensive and react emotionally. So even though

we work together as English teachers, we seem to have some sort of

protective barrier by ourselves in this sense. (Teacher C)

The excerpt 4.23 indicates implicit tension between teachers which is based on KETs' anxiety that when they reveal their linguistic uncertainty, they might be perceived as incompetent in language knowledge. The naming of "disqualified English teacher" based on linguistic abilities is also implied in the theme from open-ended responses. According to the responses, KETs experience constant comparison of English proficiency between teachers in the process of their teaching and assessment. Excerpt 4.19 has already shown that the teacher felt a lack of linguistic proficiency in class due to the comparison with a co-teaching NEST. Also, comments from Teacher D

insinuate such a view in excerpt 4.24 below:

Excerpt 4.24. I think expertise in language itself seems to be the most important competence that an English teacher should have. Especially, the expertise in judging linguistic accuracy. For example, when assessing a written sentence, sometimes judgment of its grammaticality or correctness varies between teachers. I assume those kinds of ability should be basic requirements for teachers, but it seems to differ between teachers, so sometimes it was hard for me to settle an acceptable answer when I assess students with other teachers. (Teacher D)

The excerpt 4.23 and 4.24 illustrate similar situations from a different point of view. In excerpt 4.24, Teacher D mentions an example of facing conflict due to a self-defensive reaction of another teacher in the process of assessment. In the excerpt, the teacher assumes that such conflict occurs because of "disqualified" teachers who have not fulfilled "basic requirements for teachers." Both situations shared by teacher interviewees reveal the imagined notion of "disqualified English teachers" which is judged by KETs' language proficiency and linguistic anxiety they have. Thus both

intrapersonal identity conflict and interpersonal comparison of teacher qualification seem to arise from the language learner identity that KETs still possess because of their self-recognition as non-native English speakers.

The macro level of discourse also influences enhancing the NNEST identity of KETs. The analysis of open-ended responses has shown that balancing between reliability and validity of the assessment appeared as the most concerned and problematic area. In the analysis of the interview, both issues of reliability and validity arise again at the institutional level of discourse. In the in-depth interview, the sources of those issues are revealed from a macro perspective: required test objectivity due to the college entrance system.

According to analysis, securing "objectivity of test results" is repeatedly mentioned by many interviewees and appeared as a major theme related to the test reliability that KETs concern the most when they assess students. Many difficulties that teachers face during the assessment practice are based on the implicit demand that assessment results should be objective and clear. Due to this expected objectivity, teachers were concerned that language assessment results cannot be clearly proved because language ability cannot be easily measured numerically as a matter of right and wrong. Hence, to make the grading to be supported by evidential background, teachers focus on developing precise rubrics and strengthening the evidential

basis of scores. For instance, KETs are led to stick to the concrete rubrics with more focus on the domains of linguistic accuracy to guarantee the objective basis for the scoring. This scoring tendency implies that rubrics KETs use are still highly related to the native-speaker norm.

Teacher E argues in excerpt 4.25 below that due to the relation to the college entrance system, various stakeholders including students and parents are concerned about the objectivity of assessment results.

Excerpt 4.25. I feel least confident in assessing the speaking and writing skills of students. In middle schools and high schools, all evaluations are directly related to the college entrance exam, so objectivity in assessments' results is required. In terms of objectivity, however, those language skills cannot be easily measured by numerical scales with clear proof, so I had difficulty in assessing those areas. Even though various assessment rubrics are developed, still it is hard to assess with those rubrics because it is not easy to clearly define the boundary of scoring scales when teachers actually assess students' language behavior. In the assessment, many students' language behaviors cross the boundary of scoring scales in rubrics, so teachers' subjective judgment has no choice but to intervene in

which side to put them in. It seems that students and parents are most

concerned about that. (Teacher E)

Teacher E pointed out that the college entrance system significantly influenced teachers' assessment practice. Moreover, such avoidance of KETs' "subjective" judgments implies that their expertise in assessing language behavior is questioned by other stakeholders, like students or parents. Teacher B illustrated her experience of co-teaching with a NEST and her experience of such distrust on KETs' assessment decisions as below:

Excerpt 4.26. KETs' teaching method and technique, based on my

experience, are much better than NESTs. We, Korean English teachers,

know better about students' needs and are well trained. But, when it

comes to assessments, the intuition of native speakers comes forward.

I admit that my intuition could be less dependable than NEST's one,

but I think students or parents do not seem to trust our decisions in

assessments, especially in speaking assessments. It is hard to say

whether the source of distrust comes from KETs' self-doubt or other

peoples' point of view, but I guess because of the distrust, KETs

become more dependent on the NESTs opinions when they work with

NESTs. It seems to me that native speakers' intuition and their

standards work like a shield in schools with NESTs. (Teacher B)

According to excerpt 4.25 and 4.26, KETs' assessment practice and their decisions in language assessments are questioned in terms of objectivity and are doubted by the standard of native speaker model. In the same vein, Teacher D shares her experience of using rubrics leaning toward linguistic accuracy due to the pressure of securing the objectivity of grading, while she was concerned about the validity of the test in excerpt 4.27.

Excerpt 4.27. For example, when I did assessments in writing skills, I

doubted whether the assessment was valid because the scoring

criteria were mainly about concrete aspects of the written product.

For example, the scoring was based on whether the grammatical

errors were made or not, or the number of the written sentences in the

essay, rather than scoring the organization or the content of the essay.

It was not the way I wanted to do the writing assessment, but I work

in a big school and other teachers wanted to use those kinds of

criteria, so I had no choice. (Teacher D)

In sum, KETs' identity in discourse is influenced by various factors at

the intrapersonal, interpersonal, and institutional levels of discourse and it reveals the reason why KETs' identity as the expert cannot be easily established. KETs expressed identity conflicts at the intrapersonal level as they identify themselves both as a teacher with expertise and a language learner with linguistic uncertainty. This contrasting but coexisting identity of KETs as teaching experts and language learners is also implied at the interpersonal level of discourse: the "defensive reaction" of KETs in assessment practice which implies KETs' constant anxiety as incompetent language users and the naming of "disqualified English teacher," which stigmatizes English teachers with lower language proficiency. Furthermore, at the institutional level, the sociopolitical discourse of securing "objectivity" of test results forces KETs to minimize their decisive intervention in language assessment and makes them cling to the paradigm of the native speaker model in their assessment practice. The discourse which gives priority to the fairness of the test negatively influences KETs' identity as assessors. Hence, KETs' identity construction in discourse is hindering them from recognizing and developing themselves as professional assessors.

4.4.2. Teacher Identity in Practice

Varghese et al. (2005) described identity-in-practice as an essential

constituent of teacher identity in the understanding identity. In this action-oriented approach, identity is constructed through practices and tasks and formed as a social matter. In the same vein, Wenger (1999) conceptualized three modes of belonging in identity construction through doing: engagement, imagination, and alignment.

Through engagement, teachers establish and maintain relations with others in a discourse community; they share, learn and develop the conventions and practices of the community and it is essential for constructing professional identities. According to the interviews, many KETs have already known the importance of engaging in the teachers' community to share and learn through each other's experiences for developing their professional identities. The majority of interviewees noted that discussing with other teachers was the best way to handle difficulties of assessment they face in class. Furthermore, teacher training that involve sharing teachers' experiences and practical teaching and assessment methods were emphasized as the most helpful aid in improving their self-efficacy in assessment. In-class teaching and assessment experience was mostly valued for becoming an expert teacher. It was also implied in the comments of some interviewees who were reluctant to identify themselves as language teaching experts. In the following excerpt 4.28, Teacher F found the reason for her lack of confidence from her inexperience in teaching.

Excerpt 4.28. I do not feel confident in teaching and assessing English yet. I think I learned a lot about English education so far, but I still do not connect the knowledge I have to the actual teaching or assessments. I assume that the expertise comes from theoretical background and experiences. But, I was not able to relate my knowledge of English education to what I do as a teacher. And I still do not have much experience as a novice teacher. So, I cannot think that I am an expert in English education yet. (Teacher F)

In identity construction, imagination indicates that individuals move beyond their engagement of a community in the real world and create images of the imagined communities. Through imagination, individuals legitimize their positions in the imagined communities and envision the ideal representation of the language teacher in relation to their work. In terms of imagination, KETs seem to identify themselves as EFL teachers in Korea. However, they try to differentiate their practices from "Korean-style" English teaching, as the term "Korean-style" has a negative connotation. Their endeavor to differentiate their practice from so-called Korean-style English teaching may imply that they still were not able to position themselves legitimate users of English. In other words, KETs were not able to consider themselves as legitimate English users in imagined communities because of

their non-nativeness.

When KETs illustrate their assessments, they called "Korean-style tests" and "memorization tests" to indicate assessments that were not performed properly. These terms also appeared in open-ended question responses as a pitfall of assessment practice in EFL situations that teachers continuously encounter. The negative implication of these terms indicates that KETs identify the evaluation methods previously implemented in English education in Korea as illegitimate and invalid. The excerpt shared by teacher G illustrates that KETs' self-awareness of NNEST is distinguished in their concern with "Korean-style tests" as below:

> *Excerpt 4.29. Even though I've been into English education long time but still I am not a native speaker, so when I assess speaking or writing skills I was worried if I focus too much on grammatically-oriented scoring or I was too biased toward Korean-style tests as I try to differentiate students' grades. (Teacher G)*

Excerpt 4.29 demonstrates that Teacher G considers her status as NNEST might result in conducting assessments in "Korean-style". The implied negative evaluation of "Korean-style" indicates that KETs are trying

to distinguish their identity positions from teachers who teach and assess with traditional grammar-translation method by engaging in assessments associated with "assessment for learning".

Along with "Korean-style" tests, KETs also mentioned problems with "memorization tests". According to the interviews, in the in-class speaking and writing assessment, students are inclined to memorize their scripts or drafts in advance to get better grades. The teachers also reported that some students do not have a choice but to memorize because of their low level of language proficiency. KETs argued that the "memorization test" is problematic due to its low test validity but somewhat inevitable for lower-level EFL students. They also claimed that the problem occurs because of poorly structured, product-oriented evaluations.

KETs were fully aware of the limitations of EFL circumstances, and they pointed those out as "Korean-style" and "memorization tests." Hence they focused on developing assessments for learning, which aligns with the assessment framework of education policy as well. The majority of interviewees pointed out that they prioritized what students are going to experience in the process of assessments. A few teachers shared their assessment experience and pointed out that the needs of EFL learners are different from that of ESL or native speakers; hence they argued that assessments should be considered as providing an opportunity for students to

actually use language and learn through the process. Teacher E shared her experience in the excerpt below:

> *Excerpt 4.30. The process-based writing assessment was the best one I have ever done. Because most students are depending on the use of language translators these days, when writing assessments are noticed with certain topics, they prepare for the assessments by writing a draft using a language translator and memorizing it. It means that the only effort students make is memorizing the prepared draft beforehand. I thought that it is not the objective of the evaluation. As far as I know, performance assessments including writing assessments are supposed to be an assessment for learning, not the assessment of learning. So rather than assessing what students have learned in class, now teachers should lead students to learn something through assessments. So I used the process-based writing assessment and I think it was effective for learners to learn how to write step by step. (Teacher E)*

As excerpt 4.30 illustrated, process-based assessments are accepted as an alternative and more adequate type of assessments for EFL learners as it

provides an opportunity for learners to use language and learn strategic aspects of language learning. By adjusting their assessment methods to fit the needs of EFL students, KETs identify themselves as professional EFL teachers and differentiate themselves from the negative NNEST identity.

It was also revealed that KETs change their assessment practice constantly to align with their school circumstances. KETs reported that they think about effective teaching and assessment methods despite differences in level among students and adjust tests and rubrics to fit the proficiency level of students they are currently teaching. Moreover, teachers tend to value providing an adequate level of assessment highly so students can have mastery experiences to feel a sense of accomplishment in language learning. Teacher H shared her experience of considering students' needs in planning assessments in excerpt 4.31:

Excerpt 4.31. I happen to understand that some students, who do not care for studying, will not need to learn English at all in this EFL context of Korea. I teach in a school with low-performing students and I came to realize that for some students, learning English could be meaningless, because we do not use English in our daily lives. Because of this realization, I changed my assessments. For example,

when I assess speaking abilities, I used to prefer conversational tasks such as talking in pairs and did not even consider presentations as speaking tasks due to its unidirectionality. But now I have students to do the presentation for the speaking assessments. Because, when I think of these students' circumstances, the possibility for them to use English might not involve conversational situations. But I thought that they might happen to do presentations in English someday. They are going to have a job anyway and maybe somehow they would do some business meetings. So I thought that for their lives, teaching how to do a presentation in English might be more meaningful than teaching conversations. This realization made me be aware of the significance of EFL context, and led me to think over what kind of tasks and English use can be meaningful for students in EFL context. (Teacher H)

In sum, teacher agency as an assessor in practice was revealed in terms of evaluation, imagination, and alignment. The contextual and circumstantial factors which hinder them from improving their assessment practice were also unveiled. In the assessment practice, KETs established their positions as professional language teachers so they associated their practices and activities with significant and also practical techniques that can be shared for professional development within their community. Furthermore,

many KETs were aware of their capability to exercise control over teaching and assessment practice, so they actively reorganized their assessments in order to promote a positive influence on students' language learning process.

On the other hand, contextual factors seem to negatively influence KETs identity. While KETs identified themselves as members of the EFL education community, they were not able to identify themselves as legitimate language users. They tend to consider language assessment conventions of Korean English education as invalid and attempted to implement the concept of assessment for learning to differentiate their assessment practice from former ones. This tendency implies KETs still felt vulnerable as NNESTs in their assessments and it led KETs to continuously struggle to improve their assessment practice by adopting the newest methods. However, employing a new assessment method cannot fully compensate for their perceived vulnerability as NNEST assessors. If they adhere to the native-nonnative divide, they would still feel anxious about possible bias they would have as NNESTs in their assessment practice. Hence, KETs' endeavor to identify as professional language teachers through their practice is still limited as they fail to legitimize themselves as genuine language users and reject the divisive ideology of native and non-nativeness.

CHAPTER 5.

DISCUSSION

In this chapter, the key research findings presented in Chapter 4 are summarized and discussed with regard to research questions. Section 5.1 summarizes the analysis of KETs' self-efficacy of LAL in different language skills and emerging issues related to assessment practice. Section 5.2 discusses the relation of KETs' LAL efficacy and teacher identity construction.

5.1 KETs' Self-efficacy of LAL

The purpose of the present study was to explore KETs' LAL self-efficacy beliefs and the sources that influence their assessment practice and to identify the relation of efficacy beliefs and teacher identity. Accordingly, the first research question was to explore KETs' self-efficacy beliefs in assessing different language skills. The quantitative results presented that statistically, KETs LAL self-efficacy did not show significant differences in assessing different language skills, despite the mean score differences.

Regardless of the statistical result, in qualitative analysis, many KET participants expressed LAL deficiency in assessing speaking and writing skills more than other skills and it is in line with previous studies with ESL teachers in which teachers reported that one of the most urgent areas in need of training was assessing productive skills (Hasselgreen et al., 2004). In order to reveal the status quo of KETs' LAL self-efficacy and areas of deficiency, open-ended question responses from the survey and follow-up in-depth interviews were qualitatively analyzed. Several themes were revealed as a major concern for the language assessment practice of teachers, which indicates the areas of shortcomings in their self-efficacy of LAL.

Maintaining both the validity and the reliability of the assessments notably stands out as the most complicated part of assessments for KETs. They were aware that those two domains are interrelated and need to be balanced properly in order for a test to be effective (Jeon & Oh, 2006). When KETs assess in schools, however, developing valid and reliable tests and scoring criteria was not easy for them, because when the test reliability was secured with concrete and definite scoring criteria, indefinite and abstract domains of language ability such as fluency or context are likely to be overlooked and minimized in the scoring criteria. This would result in atrophy of test validity and vice versa. In the same vein, KETs were concerned that the priority on the objectivity of scoring would lower the

validity of assessments. This tendency makes KETs focus more on the domains of linguistic accuracy, which is highly related to the native-speaker norm, so they can have the evidential basis of scores and maintain the test objectivity. This finding was supported by Kang and Lee (2012) who found that in assessments, KETs still took a reserved position in accepting the features of ELF and withheld the legitimacy of language features based on exonormative language rules. Therefore, KETs reported problems in the imbalance of the validity and reliability in their assessment practice and felt inefficient in these aspects of LAL.

Securing the practicality of the test despite a large number of students, and the polarized level differences between students also appeared as problems teachers encounter in their assessment practice. Also, the analysis demonstrated the tendency of KETs in assessing productive skills. In speaking and writing assessments, KETs tend to be unsure of scoring decisions and expressed difficulties in judging the acceptability of language use and to what extent errors should be tolerated. Many teachers also claimed that testing productive skills are degenerated into "memorization tests" with low test validity. Moreover, especially in speaking skills, some KETs revealed a lack of self-confidence not only as assessors but also as language users. Such deficiency in LAL self-efficacy is related to their negative NNEST identity, which will be elaborated on in the next section.

5.2 KETs' LAL Self-efficacy and Teacher Identity

Regarding the second research question, the relation of KETs' LAL self-efficacy and their identity as expert teachers was investigated. The qualitative results of the survey indicated a meaningful correlation between teachers' LAL efficacy and teacher identity. The Pearson correlation coefficient of the LAL self-efficacy and teacher identity has shown a positive correlation between those two, with the strength of medium (r>.40) to large (r>.50) (Plonsky & Oswald, 2014). This statistical analysis illustrates that the higher the KETs self-efficacy are the more positive identity teachers possess as professional language teachers.

Among the four language skills, LAL self-efficacy in assessing writing skills has shown the strongest correlation with the teacher identity. It corresponds to the findings of qualitative analysis of follow-up interviews in that accumulation of hand-on experiences of in-class assessments are acknowledged as valuable essence of teacher expertise. According to the analysis of interviews, KETs attempt to employ new methods relatively more in assessing writing skills and therefore they are likely to feel more efficient in assessing writing than any other skills as they become more experienced and identify themselves as expert teachers.

The identification of KETs as NNEST and their relation to teacher

LAL self-efficacy were investigated qualitatively based on a thematic analysis of semi-structured interviews. For systemic analysis of KETs' identity, the framework of the teacher identity investigation by Trent (2015) was used, and influential themes in teachers' identity construction as assessors were found as shown in figure 5.1.

Figure 5.1 Identified themes in KETs identity construction
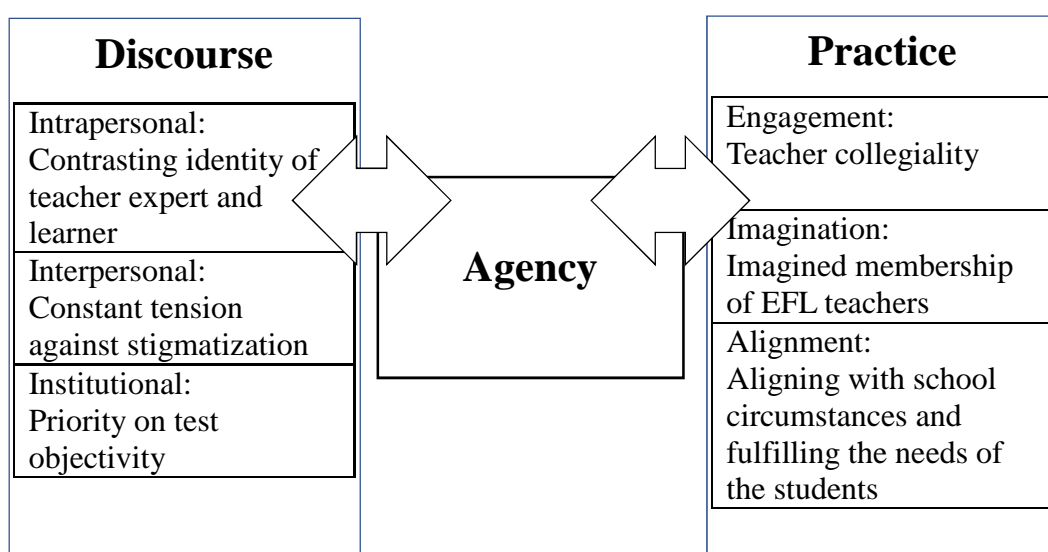
KET's identity as NNEST was revealed in both constituents of identity construction, and it negatively influenced their LAL self-efficacy. Firstly, in discourse, it is revealed that KETs' sense of LAL self-efficacy was greatly influenced by identity conflict as professional language teachers and language learners at the intrapersonal level. While they self-evaluate

themselves as experts of language education based on their knowledge of language teaching and teaching experience, they still identified themselves as language learners. This strong learner identity is based on the anxiety of a non-native speaker who constantly self-doubts his/her own language use. Hong (2013) also claimed that the strong learner identity of KET is a source of the constant tension and sense of inferiority in language proficiency. This learner identity hinders KETs from positioning themselves as convinced teaching experts and legitimate language users. The NNEST identity as a permanent language learner was also criticized by Brutt-Griffler and Samimy (2001) as they argued that the notion of 'native-speaker' is constructed by hegemonic discourses and works as bases of disempowerment.

Moreover, the stigmatization of KETs was unveiled in interpersonal tension between students and teachers and between co-working teachers. KETs believe that they are likely to be stigmatized as disqualified teachers if they are considered to have low language proficiency; hence they demonstrated a tendency of avoiding linguistic mistakes in class. It is supported by Hong's (2014) claim that expected and presumed language fluency for English teachers is the source of such constant tension of KETs which would result in enhancing their learner identity. The presupposition of teacher language fluency comes from the social ideology which considers an English teacher's competence is best judged by whether it reached a native

speaker proficiency level or not (Jenkins, 2005). The naming of disqualification is not only feared in the teacher-student relationships but also among teachers. Accordingly, KETs felt the obsessive pressure of developing their own language ability and tended to react with a defensive attitude when their scoring decision is questioned, which would cause difficulties in adjusting assessment criteria with other co-working teachers. It is consistent with Choe's (2008) and Lim's (2011) findings which mentioned KETs' continuous desire for increasing language proficiency. Hong (2014) noted that this interpersonal tension differs according to the school circumstances, especially the English level of the students they teach. In the schools with students of high language proficiency, KETs are likely to experience a sense of shame and identify as "an incompetent teacher who lacked language ability", while they identify as "a proficient language teacher" in the schools with low language proficiency students.

The teacher identity-in-discourse is also influenced by the institutional level of discourse which is interrelated with the national college admission policy. Since college-admission policy involves school record-focused selection, securing the objectivity of assessment was prioritized in schools to avoid complaint reports on the assessment by the stakeholders. These complaint reports on grades by students or parents are proven to be a critical source of teacher frustration; hence teachers are compelled to retain

an unbiased, objective explanation of their scoring (Wi, 2020). In the same vein, with the focus on test reliability, the native speaker model intervenes as a genuine standard in the pursuit of developing precise and concrete rubrics. Accordingly, KETs' discretionary power in assessments is compromised, and as an NNEST their scoring decisions tend to be distrusted by the stakeholders without the support of scoring evidence. This tendency was also found in the study of Azizjon and Ma (2019) and Sung and Jo (2015) that KETs are required to provide crystal clear parameters of assessments, because of stakeholders' distrust of school assessments, hence assessments without such criteria are avoided. This macro perspective of discourse based on the college admission policy significantly disgraces KETs' identity as professional language assessors and constrains their discretion.

KETs' identity construction in terms of practice was also found to be influenced by NNEST identity. In the study, KETs associated their practice and activities with expert identities and characterized them as practical techniques which can be developed by first-hand experience. Accordingly, KETs tend to value teacher collegiality in schools, and also they stressed the benefit of sharing such experiences of practical teaching and assessment activities within teacher communities. Also, teachers adjusted their assessments in order to fit the exact needs of the students they are teaching and suggested the implementation of process-based assessment as they aim to

realize assessment for learning. In this aspect, teacher agency was prominent and KETs seemed to be able to manage the process of teaching and assessments.

KETs were aware of their EFL context and strived to deal with contextual shortcomings. However, in their practice, KETs were not capable of identifying themselves as legitimate EIL language users. Instead, they focused on differentiating their practice from former assessment practice which was considered invalid with the naming of "Korean-style" tests, and developing effective assessment methods. This negation of Korean-style English tests indicates that KETs are still measuring students' language behavior within the ideological model of native speakers in their scoring decisions and do not consider the EIL language model as legitimate. It is in line with Nguyen's (2017) research in which NNEST teachers who appreciate the EIL perspective still appeared to identify strongly with the native speaker model in person and believed in the necessity of native-likeness. In the study, Nguyen (2017) argued that NNEST teachers could fall back into the native-speaker fallacy due to this belief, regardless of their association with EIL and multilingualism. Kang and Lee (2012) also found that in assessments, KETs were inclined to the native speaker norms and were reluctant to accept the features of ELF. Therefore, KETs still seems to be influenced by the paradigm of the native speaker model and they also neglected to seek an

alternative model of linguistic proficiency for EFL learners. In this respect, KETs are at risk of being preoccupied with negative NNEST identity, even though they struggle to identify themselves as professional language teachers.

Although each constituent of identity construction were classified into two categories as shown in Figure 5.1, it should be recognized that the elements in the diagram are interconnected and are influenced by each other (Trent, 2015). Hence, KETs' identity construction should be understood as a multidimensional and multifaceted process that is influenced by various internal and external factors. In terms of discourse and practice, the fragile aspect of KETs' identity as NNEST and its influence on their assessment practice was unveiled. KETs are continuously struggling to improve their LAL and expertise by working on their linguistic proficiency and assessment methods and expressed an urgent need for proper and practical teacher training to enhance teacher LAL (Chung & Nam, 2018; Jeon & Oh, 2006). Within the ideology of the native and non-native divide, however, KETs will not be able to identify themselves as eligible language assessors. Therefore, it is important to raise their awareness of power relations of language norms so they can get rid of implicit self-doubt and strengthen their self-confidence so they can identify themselves as legitimate speakers of English.

CHAPTER 6.

CONCLUSION

This study explored Korean English teachers' LAL efficacy and its relation to their identity as non-native professionals. Chapter 6 summarizes the major findings and pedagogical implications of the present study. Section 6.1 presents a summary of key findings and discusses its implication. Section 6.2 concludes the chapter with the limitations of the study and suggestions for further research.

6.1 Findings and Implications

The primary aim of this study was to investigate the efficacy of language assessment literacy (LAL) of Korean English teachers (KET) in secondary schools and perceived areas of deficiencies in their assessment practices. The second objective was to investigate their identity as non-native English speaking teachers (NNESTs) and its relation to their LAL self-efficacy belief. A mixed-method research design with a survey and an interview were adopted to collect and analyze data both in a quantitative and

qualitative ways. The survey was conducted to reveal a statistical relation of KETs' LAL self-efficacy and the teacher identity, and semi-structured interviews were performed to obtain specific data about contextual aspects of their LAL self-efficacy and teacher identity formation.

The present study results confirmed that KETs' identity includes negative features of NNEST identity and those features are found to be influential to their self-efficacy in language assessment literacy. The findings of the present study suggest some pedagogical implications.

First, in-service teacher training programs to support KETs to disengage from native-speaker norms and identify as non-native professionals should be provided to many teachers. The findings of this study proved that the negative self-perception of NNEST impedes KETs' teaching and assessment ability in terms of the discourse and the practice. Gaining knowledge of the EIL perspective and knowing that appreciation of diverse English varieties is supported by empirical research have been found to be empowering NNESTs to see beyond the dichotomy of native-ness and eventually to become critical practitioners (Nguyen, 2017). Since the 2015 Revised National English Curriculum has outweighed performance assessments, KETs are required to practice diverse forms of alternative assessments. Teachers, to some extent, can reap beneficial gains by training for strategies and methods and enhancing their knowledge base of

assessments. However, to relate their knowledge and transform the assessment practice, teachers' interpretive frameworks need to be improved and reconceptualization of the current negative NNEST identity of KETs is essential.

Second, KETs' collegiality should be strongly supported within schools in order for teachers to reinforce their identity as professional language teachers. In the study, KETs considered sharing their teaching and assessment practices with other English teachers as the most effective way to deal with problematic situations they encounter in schools. Moreover, they believed that feasible and practical strategies of teaching and assessment can be learned through in-hand experience, and valued sharing those tips and experiences of other KETs. Teacher collegiality fosters a positive atmosphere where teachers can actualize their teacher agency. Also, collegial supervision with an insider's view supports them to be more efficient in their assessment practice. Hence, teacher collegiality would encourage KETs to construct professional language teacher identities both in terms of discourse and practice. Considering that teachers' efficacy beliefs undergo changes and their assessment efficacy is strengthened and weakened continuously based on the influence of context, promoting teacher collegiality should be taken into account in developing KETs' LAL efficacy.

6.2 Limitations and Suggestions for Further Research

There are several limitations in the current study. Firstly, due to its small sample size, its results lack generalizability in confirming influential sources of KETs' LAL self-efficacy and identity construction which are contextually situated. The researcher explored teachers' language assessment practice through mixed-method approach, but the applicability of the study results is yet limited. With a larger sample, more applicable findings of teacher identity construction will be discovered.

Secondly, the survey needs to be made to minimize carry-over effects. The carry-over effect indicates that the survey responses can be affected by prior items. However, in this study, the sections of Part A were given in the same order to all participants, and it could have affected their answers. Hence, changing the order of the survey sections for each participant should be considered in order to gain precise results by minimizing the carry-over effect.

Lastly, the research employed methods of an online survey and interviews which are both depending on self-reported data. This indicates that other efficacy-related information could be missed out on in the data collection. Therefore, further research is suggested with research designs in multiple respects by including classroom observations and a collection of teachers' reflective journals. Despite these limitations, the study addresses

significant implications in promoting the professional growth of KETs and empowering them.

REFERENCES

American Council on the Teaching of Foreign Languages. (2012). *ACTFL proficiency guidelines*. Retrieved from http://www.actfl.org/publications /guidelinesand-manuals/actfl-proficiency-guidelines-2012/english

Aneja, G. A. (2016). (Non)native speakered: Rethinking (non)nativeness and teacher identity in TESOL teacher education. *TESOL Quarterly, 50*(3), 572-596.

Azizjon, K., & Ma, J. H. (2019). Korean English teachers level of assessment literacy and their perceptions of school assessments. *Studies in Linguistics, 52*, 285-303

Bae, Y. (2015). The influence of native-speakerism on CLIL teachers in Korea. In A. Holliday, P. Aboshiha & A. Swan (Eds.), *(En)Countering native-speakerism: Global perspectives* (pp. 75-90). London: Palgrave Macmillan.

Bae, W., & Shin, S. (2016). An exploratory study of status quo of standards-based testing system and its impacts on English language instruction in Secondary Schools. *Secondary Education Research, 64*(3), 611-634.

Bang, I., & Chon, S. (2011). An analysis of high school English curriculum implementation by the high school record rating system. *CNU Journal of Educational Studies, 32*(1), 1-20.

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice, 5*(1), 7-74.

Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2004). Working inside the black box: Assessment for learning in the classroom. *Phi Delta Kappan, 86*(1), 8-21.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*, 77–101.

Brutt-Griffler, J., & Samimy, K. K. (2001). Transcending the nativeness paradigm. *World Englishes*, *20*(1), 99-106.

Canagarajah, A. S. (2004). Subversive identities, pedagogical safe houses, and critical learning. In B. Norton & K. Toohey (Eds.), *Critical pedagogies and language learning* (pp. 116–137). New York, NY: Cambridge University Press.

Canagarajah, A. S. (2013). Negotiating translingual literacy: An enactment. *Research in the Teaching of English*, *48*(1), 40-67.

Chung, S. J., & Nam, Y. (2018). Language assessment literacy of Korean EFL teachers: An investigation of their training experiences and needs. *Modern English Education, 19*(1), 38-48.

Choe, H. S. (2008). Korean nonnative English speaking teachers' identity construction in the NS-NNS dichotomy. *English Language and Linguistics*, *26*, 109-127.

Creswell, J. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches* (4th ed.). Thousand Oaks, CA: SAGE.

Cook, V. (2016). Where is the native speaker now?. *TESOL Quarterly, 50*(1), 186-189.

Davies, A. (2008). Textbook trends in teaching language testing. *Language Testing*, *25*(3), 327-347.

DeLuca, C., & Klinger, D. A. (2010). Assessment literacy development: Identifying gaps in teacher candidates' learning. *Assessment in Education: Principles, Policy & Practice, 17*, 419–438.

Dornyei, Z. (2007). Qualitative, quantitative and mixed methods research. In *Research methods in applied linguistics* (pp. 24-47). Oxford: Oxford University Press.

Erickson, G. (2020). Finding out what learners know-and…? Reflections on teachers' language assessment literacy. In D. Tsagari (Ed.). *Language Assessment Literacy: From Theory to Practice* (pp. 29-49). Newcastle, UK: Cambridge Scholars Publishing.

Fulcher, G. (2012). Assessment literacy for the language classroom. *Language Assessment Quarterly, 9*(2), 113-132.

Ha, M. A. & Min, C. K. (2008). The constituents of English teachers' professional competency in EFL context. *Journal of the Korea English Education Society, 7*(2), 87-110.

Hasselgreen, A., Carlsen, C., & Helness, H. (2004). *European survey of language testing and assessment needs. Part one: General findings*. Retrieved from http://www.ealta.eu.org/resources.htm

Hill, K., & McNamara, T. (2012). Developing a comprehensive, empirically based research framework for classroom-based assessment. *Language Testing, 29*(3), 395-420.

Holliday, A. (2005). *The struggle to teach English as an international language.* New York: Oxford University Press.

Holliday, A. (2006). Native-speakerism. *ELT Journal*, *60*(4), 385-387.

Hong, Y. S. (2013). Living as non-native English teachers in Korean elementary school contexts: Focusing on teacher identity formation. *English Language & Literature Teaching, 19*(4), 427-453.

Hwang, E. & Lee, H. (2018). A systematic review on the identity of Korean English teachers as non-native English speakers. *The Journal of Foreign Studies, 46*, 99-128.

Inbar-Lourie, O. (2008). Constructing a language assessment knowledge base: A focus on language assessment courses. *Language Testing, 25*(3), 385-402.

Jenkins, J. (2005) Implementing an international approach to English pronunciation: The role of teacher attitudes and identity. *TESOL Quarterly, 39*(3): 535–543.

Jenkins, J. (2009). Exploring attitudes towards English as a lingua franca in the East Asian context. In K. Murata & J. Jenkins (Eds.), *Global Englishes in Asian contexts* (pp. 40-56). London: Palgrave Macmillan

Jeon, B. M. & Oh, J. I. (2006). Secondary school English teachers' perceptions of their competence in student assessment. *Studies in English Education, 11*(2), 19-43.

Jeong, H. (2013). Defining assessment literacy: Is it different for language testers and non-language testers?. *Language Testing, 30*(3), 345-362.

Kang, Y. S., Jin, S. H., & Jang, S. Y. (2015). Exploring secondary school English teachers' perceptions and management of constructed response assessment as professional testers. *Studies in English Education, 20*(2), 323-353.

Kang, S. & Lee, S. (2012). A survey of Korean English teachers perceptions and attitudes about the common features of English as a lingua franca: Focusing on the lexical and grammatical features. *Korea Journal of English Language and Linguistics 12*(3), 379-401.

Kim, H. K. (2013). "English is power": Narratives of Asian English teachers in language education. *English Language & Literature Teaching, 19*(1), 89-109.

Kim, Y. M. & Cheong, S. H. (2012). The effect of collaborative teacher development meeting on constructing teacher identity and developing a college English curriculum. *Foreign Languages Education, 19*(4), 399-424.

Kwon, B. Y. (2013). An analysis of Korean English teachers' motivation for the teaching profession and the degree of self-efficacy: Comparison with Korean language and math teachers. *Journal of Holistic Education, 17*(1), 1-24.

Lee, H. (2010). Analyzing the effects of collaborative action research from the teacher identity perspective. *English Teaching*, *65*(1), 161-188.

Lee, M. (2018). A study of high school English teachers' teaching English speaking and performance assessment. *Journal of the Korea English Education Society, 17*(1), 107-126.

Leung, C. (2005). Convivial communication: Recontextualizing communicative competence. *International Journal of Applied Linguistics*, *15*(2), 119-144.

Leung, C. (2007). Dynamic assessment: Assessment for and as teaching? *Language Assessment Quarterly, 4*(3), 257–278.

Malone, M. E. (2013). The essentials of assessment literacy: Contrasts between testers and users. *Language Testing, 30*(3), 329-344.

McNamara, T. F. (1996). *Measuring second language performance.* New York: Longman Publishing Group.

McNamara, T. F., & Roever, C. (2006). *Language testing : The social dimension*. Malden, MA: Blackwell.

Medgyes, P. (1992). Native or non-native: Who's worth more?. *ELT journal*, *46*(4), 340-349.

Monfared, A. & Khatib, M. (2018). English or Englishes?: Outer and expanding circle teachers' awareness of and attitudes towards their own variants of English in ESL/EFL teaching contexts. *Australian Journal of Teacher Education, 43*(2), 56-75.

Newman, K. L. (2005). *What difference does it make? Nonnative English speaking teachers in K–12 school settings*. Doctoral dissertation, Indiana University, Bloomington, IN.

Nguyen, M. X. N. C. (2017). TESOL teachers' engagement with the native speaker model: How does teacher education impact on their beliefs?. *RELC Journal, 48*(1), 83-98.

Park, T. (2016). Examining how performance assessment in English is being implemented in high schools and how it is perceived by teachers and students. *Secondary English Education, 9*(3), 77-98.

Park, T., & Chang, J. (2017). A study on high school teachers' opinions about issues of communicative English teaching and assessment. *Korean Journal of English Language and Linguistics, 17*(4), 839-863.

Pennington, M. C. (2014). Teacher identity in TESOL: A frames perspective. In Y. L. Cheung, S. B. Said & K. Park (Eds.). *Advances and current trends in language teacher identity research* (Vol.4, pp. 16-30). New York, NY: Routledge.

Phillipson, R. (1992). *Linguistic imperialism.* Oxford: Oxford University Press.

Pill, J., & Harding, L. (2013). Defining the language assessment literacy gap: Evidence from a parliamentary inquiry. *Language Testing, 30*(3), 381-402.

Plonsky, L., & Oswald, F. L. (2014). How big is "big"? Interpreting effect sizes in L2 research. *Language Learning, 64*(4), 878-912.

Popham, W. J. (2011). Assessment literacy overlooked: A teacher educator's confession. *The Teacher Educator, 46*(4), 265-273.

Rea-Dickins, P. (2001). Mirror, mirror on the wall: Identifying processes of classroom assessment. *Language Testing, 18*(4), 429–462.

Rea-Dickins, P. (2006). Currents and eddies in the discourse of assessment: A learning-focused interpretation[1]. *International Journal of Applied Linguistics, 16*(2), 163–188.

Scarino, A. (2013). Language assessment literacy as self-awareness: Understanding the role of interpretation in assessment and in teacher learning. *Language Testing, 30*(3), 309-327.

Shin, D. & Park, S. (2013). A trend analysis of research papers on identity in language education: From the perspectives of post-structuralistic theory. *The Sociolinguistic Journal of Korea, 21*(1), 99-127.

Song, M. Y., & Kim, K. H. (2007). Development and validation of teacher's assessment professional competency test. *Journal of Curriculum Evaluation, 10*(1), 47-74.

Sung, M., & Jo, V. H. (2015). Sociopolitical and contextual influences on teacher-produced achievement tests of English in Korean high schools. *SNU Journal of Education Research, 24,* 115-148.

Stiggins, R. J. (1991). Assessment literacy. *Phi Delta Kappan, 72*(7), 534-539.

Taylor, L. (2009). Developing assessment literacy. *Annual Review of Applied Linguistics, 29*, 21-36.

Taylor, L. (2013). Communicating the theory, practice and principles of language testing to test stakeholders: Some reflections. *Language Testing, 30*(3), 403-412.

Trent, J. (2015). Towards a multifaceted, multidimensional framework for understanding teacher identity. In Y. L. Cheung, S. B. Said & K. Park (Eds.). *Advances and current trends in language teacher identity research* (Vol.5, pp. 44-58). New York, NY: Routledge.

Varghese, M., Morgan, B., Johnston, B., & Johnson, K. A. (2005). Theorizing language teacher identity: Three perspectives and beyond. *Journal of Language, Identity, and Education, 4*(1), 21-44.

Vogt, K., & Tsagari, D. (2014). Assessment literacy of foreign language teachers: Findings of a European study. *Language Assessment Quarterly, 11*(4), 374-402.

Wenger, E. (1999). *Communities of practice: Learning, meaning, and identity.* New York, NY: Cambridge university press.

Wi, H. J. (2020). *English teachers' self-efficacy beliefs in performance assessment: A case study of six Korean EFL high school teachers.* Unpublished master's thesis, Seoul National University, Seoul.

Yi, Y. E., & Yang, S. E. (2009). A qualitative study on the experience of mothers sending their children to English kindergarten. *Korean Journal of Human Ecology, 18*(5), 985-994.

APPENDICES

Appendix 1. Questionnaire

첫번째 설문은 선생님의 평가 수행에 관하여 조사하는 설문입니다. 평소 선생님께서 듣기, 읽기, 말하기, 쓰기의 네 영역별로 평가를 시행할 때 갖는 생각과 다음의 각 문장이 얼마나 일치하는지에 따라 설문에 응답해주시면 됩니다. 다음 척도를 활용하여 응답해 주시기 바랍니다.

```
1 = 전혀 그렇지 않다
2 = 그렇지 않다
3 = 보통이다
4 = 그렇다
5 = 매우 그렇다
```

섹션	문항		응답				
			전혀 그렇지 않다 1	그렇지 않다 2	보통이다 3	그렇다 4	매우 그렇다 5
Testing listening skills	학생들의 '듣기' 능력을 평가할 때 나는...						
	1	핵심 성취기준에 따라 학생의 성취수준을 파악하기에 적합한 듣기평가 방법을 선택할 수 있다.					
	2	평가의 목적과 내용에 부합하는 듣기평가 문항과 채점 기준을 개발할 수 있다.					
	3	평가의 타당도, 신뢰도를 이해하여 듣기평가 도구의 질을 점검하고 개선할 수 있다.					
	4	듣기 학습목표 및 평가 목표에 따라 적절한 채점 기준을 설정할 수 있다.					
	5	듣기평가 채점기준에 근거하여 정확한 채점을 실시할 수 있다.					

	6	수집한 듣기 평가 결과를 통해 학생의 학습 성취 수준에 대한 종합적 판단을 내릴 수 있다.					
	7	듣기 평가 실시 후 평가 결과를 바탕으로 학생들에게 피드백을 적절하게 제공할 수 있다.					
	8	듣기 평가 실시 후 평가활동을 반성하여 추후 듣기 수업의 내용, 수준, 교수자료 등과 평가방법을 개선할 수 있다.					
	9	듣기 평가 실시에서 어려웠던 점은 무엇이 있었는지 자유롭게 적어주세요. (한 가지 이상 기재 가능)	개방형 문항				
Testing reading skills		학생들의 '읽기' 능력을 평가할 때 나는...					
	10	핵심 성취기준에 따라 학생의 성취수준을 파악하기에 적합한 읽기평가 방법을 선택할 수 있다.					
	11	평가의 목적과 내용에 부합하는 읽기평가 문항과 채점 기준을 개발할 수 있다.					
	12	평가의 타당도, 신뢰도를 이해하여 읽기평가 도구의 질을 점검하고 개선할 수 있다.					
	13	읽기 학습목표 및 평가 목표에 따라 적절한 채점 기준을 설정할 수 있다.					
	14	읽기평가 채점기준에 근거하여 정확한 채점을 실시할 수 있다.					
	15	수집한 읽기 평가 결과를 통해 학생의 학습 성취 수준에 대한 종합적 판단을 내릴 수 있다.					
	16	읽기 평가 실시 후 평가 결과를 바탕으로 학생들에게 피드백을 적절하게 제공할 수 있다.					
	17	읽기 평가 실시 후 평가활동을 반성하여 추후 읽기 수업의 내용, 수준, 교수자료 등과 평가방법을 개선할 수 있다.					
	18	읽기 평가 실시에서 어려웠던 점은 무엇이 있었는지 자유롭게 적어주세요. (한 가지 이	개방형 문항				

		상 기재 가능)					
Testing speaking skills		학생들의 '말하기' 능력을 평가할 때 나는...					
	19	핵심 성취기준에 따라 학생의 성취수준을 파악하기에 적합한 말하기평가 방법을 선택할 수 있다.					
	20	평가의 목적과 내용에 부합하는 말하기평가 문항과 채점 기준을 개발할 수 있다.					
	21	평가의 타당도, 신뢰도를 이해하여 말하기평가도구의 질을 점검하고 개선할 수 있다.					
	22	말하기 학습목표 및 평가 목표에 따라 적절한 채점 기준을 설정할 수 있다.					
	23	말하기평가 채점기준에 근거하여 수행평가에서 정확한 채점을 실시할 수 있다.					
	24	수집한 말하기 평가 결과를 통해 학생의 학습 성취 수준에 대한 종합적 판단을 내릴 수 있다.					
	25	말하기 평가 실시 후 평가 결과를 바탕으로 학생들에게 피드백을 적절하게 제공할 수 있다.					
	26	말하기 평가 실시 후 평가활동을 반성하여 추후 말하기 수업의 내용, 수준, 교수자료 등과 평가방법을 개선할 수 있다.					
	27	말하기 평가 실시에서 어려웠던 점은 무엇이 있었는지 자유롭게 적어주세요. (한 가지 이상 기재 가능)	개방형 문항				
Testing writing skills		학생들의 '쓰기' 능력을 평가할 때 나는...					
	28	핵심 성취기준에 따라 학생의 성취수준을 파악하기에 적합한 쓰기평가 방법을 선택할 수 있다.					
	29	평가의 목적과 내용에 부합하는 쓰기평가 문항과 채점 기준을 개발할 수 있다.					
	30	평가의 타당도, 신뢰도를 이해하여 쓰기평가도구의 질을 점검하고 개선할 수 있다.					

	31	쓰기 학습목표 및 평가 목표에 따라 적절한 채점 기준을 설정할 수 있다.					
	32	쓰기평가 채점기준에 근거하여 서술형, 논술형, 수행평가에서 정확한 채점을 실시할 수 있다.					
	33	수집한 쓰기 평가 결과를 통해 학생의 학습 성취 수준에 대한 종합적 판단을 내릴 수 있다.					
	34	쓰기 평가 실시 후 평가 결과를 바탕으로 학생들에게 피드백을 적절하게 제공할 수 있다.					
	35	쓰기 평가 실시 후 평가활동을 반성하여 추후 쓰기 수업의 내용, 수준, 교수자료 등과 평가방법을 개선할 수 있다.					
	36	쓰기 평가 실시에서 어려웠던 점은 무엇이 있었는지 자유롭게 적어주세요. (한 가지 이상 기재 가능)	개방형 문항				

두 번째 설문은 교사의 정체성에 관해 조사하는 설문입니다. 다음의 각 문장이 선생님께서 영어 교사로 일하며 갖는 생각과 얼마나 일치하는지에 따라서 응답을 해주시면 됩니다. 다음 척도를 활용하여 응답해 주시기 바랍니다.

1 = 전혀 그렇지 않다
2 = 그렇지 않다
3 = 보통이다
4 = 그렇다
5 = 매우 그렇다

섹션	문항	응답				
(B)		전혀 그렇지 않다	그렇지 않다 2	보통이다 3	그렇다 4	매우 그렇다 5

			1				
37	나는 잘 가르치는 영어교사이다.						
38	나의 영어 구사 능력은 영어를 잘 가르치기에 충분하다.						
39	내가 한국인 학습자로 영어를 배운 경험은 학생들을 가르치는데 도움이 된다.						
40	나는 영어를 세계어로서 이해하고 있으며 다양한 억양과 표현의 영어를 가르치는 것을 중요하게 수업에서 다룬다.						
41	나의 영어 유창성은 뛰어나다.						
42	나는 영어 발음이 좋아서 교수활동 중에 영어로 말할 때 부담을 느끼지 않는다.						
43	나는 영미문화를 잘 알고 있다.						
44	학생들 앞에서 실수하면 내가 영어를 못한다고 여겨질까봐 걱정이 된다.						
45	영어가 나의 모국어가 아니기 때문에 내 영어 능력은 부족한 부분이 있다.						
46	영어를 모국어처럼 하는 학생을 가르친다면 나의 영어능력 때문에 부담을 느낄 것이다.						
47	영어 원어민 교사와 협력수업을 한다면 원어민과 나의 영어능력차이가 신경 쓰일 것이다.						
48	나는 영어를 모국어로 하지 않기 때문에 이상적인 영어교사가 되기는 어렵다.						
49	영어 평가나 교수 과정에서 영어능력 때문에 어려웠던 점이 있다면 무엇이 있었는지 적어주세요.	개방형 문항					

APPENDIX 2: Interview Questions

1. 선생님께서 학교에서 진행했던 영어과목 평가를 생각해 보셨을 때 이때 어떤 평가가 잘 진행됐다, 혹은 평가를 잘했다 라고 생각하시는 게 있으신가요?

2. 평가하면서 어려웠거나 잘 안됐다고 생각하신 것도 있으신가요?

3. 그런 어려운 점에 어떻게 대처하셨나요?

4. 그러면 이제 수행평가에 대한 질문을 드릴게요,

5. 선생님께서 학생의 영어 능력을 평가하실 때 영어의 말하기 듣기 읽기 쓰기 중 어떤 영역을 평가하는 것에서 선생님께서 특별히 자신이 있으신가요?

6. 자신감 있게 시행하기 어려운 영역과 이유는 무엇인가요?

7. 평가나 수업에서 언제 자신감이나 동기가 가장 상실되나요?

8. 교사의 평가수행이 어려운 원인이 무엇이라고 생각하십니까?

9. 그런 어려운 점에 어떻게 대처하셨나요?

10. 선생님께서 평가 하실 때 특별히 중요하게 생각하시는 것은 무엇인가요?

11. 영어 평가의 비전문가를 1, 전문가를 10이라고 하면 선생님께서는 자신이 몇 정도라고 생각하시나요? 그렇게 생각하시는 이유는 뭐가 있으신가요?

12. 선생님께서 생각하시기에 교사들은 영어교육 전문가로서 인정받고 있다고 생각하시나요? 어떤 이유로 그렇다 혹은 그렇지 않다고 생각하시나요?

13. 선생님께서는 자신이 영어교육의 전문가라고 생각하시나요?

14. 어떤 이유로 그렇다 혹은 그렇지 않다고 생각하시나요?

15. 영어 교육의 비전문가를 1, 전문가를 10이라고 하면 선생님께서는 자신이 몇 정도라고 생각하시나요? 그렇게 생각하시는 이유는 뭐가 있으신가요?

16. 학생 수행을 평가 할 때 영어가 외국어라서 평가가 더 수월했던 적이 있나요?

17. 학생 수행을 평가 할 때 영어가 외국어라서 평가가 어려우셨던 적이 있나요?

18. 선생님께서 수업 하실 때 영어가 외국어여서 더 유리한 점이 있다면 어떤 것이 있을까요?

19. 선생님께서 수업 하실 때 영어가 외국어여서 더 불리한 점이 있다면 어떤 것들이 있을까요?

20. 선생님께서는 영어를 잘 한다는 게 어떤 거라고 생각하시나요?

21. 선생님께서 생각하시기에 좋은 영어교사는 어떤 능력이 있어야 하나요?

22. 영어 유창성이 영어 교사에게 얼마나 중요한 능력이라고 생각하시나요?

23. 선생님께서는 자신의 영어 능력이 어떻다고 생각하십니까?

24. 선생님께서는 원어민처럼 영어를 할 수 있으면 좋겠다고 생각하신 적이 있나요? 어떤 이유로 그렇게 생각하셨는지 궁금합니다.

25. 원어민 교사가 있다면 평가 과정에서 도움을 받을 만한 부분이

있다고 생각하십니까?

26. 원어민 같은 유창성이 있으면 선생님께서 평가 하시거나 가르치시는데 더 도움이 될 거라고 생각하시나요?

27. 선생님의 영어능력이 영어 수업이나 학생의 영어 능력을 평가할 때 어떤 영향력이 있었습니까?

28. 선생님께서 학습자로서 영어를 배울 때의 경험과 평가의 어려움이 연관되는 부분이 있나요?

29. 선생님께서 영어교사로서 자신감을 느끼시는 부분은 어떤 것이 있나요?

30. 선생님께서 영어교사로서 한계점을 느끼시는 부분은 어떤 것이 있나요?

31. 그런 한계점을 넘어서서 발전시키기 위해 어떤 노력을 하시나요?

32. 선생님께선 자신의 평가 효능감을 높이기 위해서는 어떤 것이 가장 도움이 된다고 생각하시나요?

33. 선생님께서 한국에서 외국어인 영어를 가르치는 교사라는 것이 선생님께 어떤 의미를 갖는지 선생님의 관점이 궁금합니다.

34. 지금까지 말씀하신 것들 중에 더 생각나시거나 더 얘기하고 싶은 부분이 있으신가요?

국 문 초 록

　　　효과적인 언어 교수학습에 있어서 평가는 중요한 역할을 수행한다. 의사소통능력 중심의 영어교육을 위해 2015 개정교육과정은 교수학습평가의 일체화를 강조하면서 과정중심평가와 수행평가의 시행을 확대하였다. 따라서 효과적이고 효율적인 평가 수행에 영향을 미치는 교사의 평가 효능감의 중요성이 나날이 강조되고 있다. 그러나 여러 연구에 따르면 한국의 영어교사들은 아직 낮은 평가효능감을 보이며 평가 전문성 개발에 어려움을 겪고 있다. 실질적으로 학교 평가의 질을 개선하고 한국 영어교사의 평가 효능감을 높이기 위해서는 한국의 영어교육이라는 맥락 안에서 형성된 교사의 정체성이 교사의 평가효능감에 미치는 영향을 면밀히 탐구하는 것이 필요하다.

　　　이 연구의 목적은 EFL 상황의 한국 중등학교 영어교사의 평가 효능감과 그들의 교사 정체성이 어떤 연관성을 가지는지를 탐색하는 것이다. 이를 통해 그들이 평가 효능감을 높이고 나아가 평가를 효과적으로 실시할 수 있도록 돕기 위한 방법을 고찰해보고자 한다.

　　　이를 위해 한국의 중등학교 영어 교사 70명이 자발적으로 온라인 설문조사에 참여하였다. 온라인 설문조사지는 49개 항목으로 이루어져 있으며 두 개의 하위 섹션으로 나뉘어 각각 언어기능별 평가 효능감과 전문가로서의 교사 정체성에 관한 질문으로 구성되었다. 설문지 응답을 분석하여 평가 효능감이 교사 정체성과 연관이 있는지를 확인하였다. 이후 9명의 교사를 무선표집하여 개별 인터뷰를 진행하였다. 이들은 자신의 평가 효능감과 교사로서의 자기 정체성에 대한 의견을 주었고, 이는 담론과 실행이라는 정체성 형성의 두 요인으로 구별되어 분석되었다.

　　　연구 결과에 따르면 많은 한국인 영어교사들은 자신을 전문가로

정체화하는 교사의 경우 높은 평가효능감을 보여 두 요인간에 유의미한 상관관계를 보였다. 교사의 정체성 형성에는 EFL이라는 한국의 영어교육 상황에 따른 비원어민 교사로서의 자기 인식이 담론과 실행의 두 영역 모두에서 영향을 미치는 것으로 드러났다. 담론적으로 한국인 영어교사들은 영어 유창성을 갖추어야 한다는 담론에 따라 비원어민 화자로서 학습자 정체성을 형성했고, 이에 교사 정체성이 대립되면서 자신을 영어교육 전문가로 정체화 하는데 어려움을 겪었다. 실행적인 면에서 교사들은 동료 의식을 가지고 교수평가의 경험적 지식을 나누는 등 전문가로서 자신을 정체화하려는 주도적인 모습을 보였다. 그러나 평가 규준에 있어서는 EFL 학습자의 필요에 맞춘 세계어로서의 영어가 아닌 원어민중심주의적 인 규준을 답습하는데 머물러 평가 전문가로서의 정체성을 형성하고 평가 효능감을 높이는데 잠재적인 방해 요소가 있음을 보여주었다.

본 연구는 연구대상이 소수라는 점과 방법론적 측면에서 일부 한계가 있었지만 교사의 전문가로서의 정체성 형성과 평가 효능감의 관계를 면밀히 살펴봄으로써 한국의 중등학교 교사들이 어떻게 평가 효능감을 높일 수 있는지에 대하여 통찰을 제시하고 이들이 보다 성공적으로 평가를 시행하는데 도움이 될 것으로 기대한다.

주요어: 교사 효능감, 평가 효능감, 교사 정체성, 언어 평가 문해력, 외국어 평가, 한국의 영어교사

학 번: 2020-28075