

# Review of Language Testing for Specific Purposes

Youn–Kyoung Lee  
(Catholic University of Daegu)

**Lee, Youn-Kyoung. (2015). Review of Language Testing for Specific Purposes. *The Linguistic Association of Korea Journal* 23(3), 31-48.** The present study aims at reviewing tests in business English which is a part of language for specific purposes (LSP) testing and related washback effects in this area of testing. In particular, the Test of English for International Communication (TOEIC) listening and reading tests, TOEIC writing tests, Oxford International Business English Certificate (OIBEC) and the Business Language Testing Service (BULATS) were analyzed in terms of degrees of test specificity and appropriateness of use in Korean business contexts. The results of the study showed that degrees of test specificity were different depending on the test even though all of the three tests are categorized into LSP testing; furthermore, these tests do not seem to be appropriate to use in Korean business contexts in order to evaluate test takers' genuine business English ability. The study finally suggested future directions in this area of testing.

**Key Words:** LSP, TOEIC, OIBEC, BULATS, washback effects

## 1. Introduction

Testing in business English is a crucial part in the field of Language for Specific Purposes (LSP), and its purpose is to assess a test-taker's second language ability in the relevant workplace. Most well-known tests in the fields of business are general performance-based, since they ask examinees to accomplish specific communicative tasks in the workplace (O'Loughlin, 2008). In particular, workplace assessments often include tasks which simulate the demands of particular real-world work related situations. The increased use of performance assessments in business contexts can be comprehended as "a global

trend to demonstrable outcomes of learning in concrete, practical and relevant skills” (McNamara, 1996, p. 36).

In Korea, business English ability is considered as an indispensable tool for success in virtually all business related situations, since Korea has heavily depended on international business for its economic growth. Interestingly, testing in business English, such as the Test of English for International Communication (TOEIC) listening and reading test plays a predominant role to assess test takers’ business English ability. Therefore, “most Korean companies require their applicants to submit a standardized English test score report and consider it an essential prerequisite for employment” (Choi, 2008, P. 40) even though the TOEIC listening and reading test does not assess test takers’ authentic business English skills.

Despite the importance of LSP testing in Korean contexts, studies in this area of testing in Korean academia are still in its infancy. Particularly, few studies seem to systematically analyze tests in English for business purposes and deal with related issues of the tests. Accordingly, the present study aims at reviewing the standardized business English tests (TOEIC, OIBEC, and BULATS) in term of degrees of test specificity and appropriateness of use in Korean business contexts. The study also focuses on analyzing washback effects of the tests and suggesting future directions in the area of the study.

## 2. Testing in Business English

### 2.1. The Test of English for International Communication (TOEIC)

In order to analyze testing in business English, it is worth exploring the different kinds of the testing of language for business purposes. A variety of standardized business English tests, developed by foreign testing institutes, are being used to measure overall business English ability in many different contexts. Some tests consist of multiple-choice formats, some are performance-based tests, and some are competency-based assessment.

In particular, the Test of English for International Communication (TOEIC) listening and reading test is one of the most well-known standardized tests of

business English, especially in Korea and Japan (Douglas, 2000), and as of 2012, it was taken by over six million candidates across more than 120 countries (Educational Testing Service: TOEIC, 2013). In Korea, 1.9 million candidates took the TOEIC listening and reading test in 2007, and the test consisted of 47.1% of all English testing for adult English learners (Choi, 2008). Furthermore, 90.99 % of the employees in major Korean companies, such as Samsung, Hyundai, LG, used the TOEIC listening and reading score as an employment exam (Park et al., 1998). This figure obviously shows the predominant role of the TOEIC in English proficiency testing in the Korean society.

In the following Table 1, the composition of the TOEIC listening and reading test is summarized (ETS, 2012, p. 2).

Table 1. Composition of the TOEIC Listening and Reading Test

Part	Content	K
Listening section		
1	Photographs	10
2	Question-response	30
3	Conversations	30
4	Talks	30
Reading section		
5	Incomplete sentences	40
6	Text completion	12
7	Reading comprehension	
	Single passages	28
	Double passages	20

*Note.* K= number of questions.

The TOEIC listening and reading test is comprised of 200 multiple-choice items divided into listening and reading sections (ETS TOEIC, 2013). The listening section consists of four subsections and is administrated by an audiocassette or compact disc for 45 minutes (ETS TOEIC, 2013; Yo & Rie, 2011). The reading section includes three subsections and lasts for 75 minutes (ETS TOEIC, 2013). In terms of reporting the test score, the number of correct responses in each section, Listening and Reading, is converted to a number on a scale of 5 to 495 (ETS TOEIC, 2013). The total score is provided by adding the scaled scores from the two sections together, and equals between 10 to 990 points (ETS TOEIC, 2013).

However, despite its business and commercial appearances, the TOEIC listening and reading test focuses on largely decontextualized aspects of language knowledge (Douglas, 2000) because the test is only composed of multiple-choice items which require test takers to choose correct answers. Furthermore, it is highly problematic to employ the test score for uses other than those intended by the test developers (Choi, 2008). Choi's (2008) study revealed that a number of Korean test takers had negative views on preparing the TOEIC listening and reading test because they had to enhance their test-taking strategies rather than their authentic English ability to improve their test score (Choi, 2008). Since the high test score on the TOEIC listening and reading test has been considered as a prerequisite for employment in the Korean society, this phenomenon and related negative washback effects have become a social issue (Jung, 2010; Kim, 2005). Considering this situation, it is not surprising that the Korean examinees have limited English ability despite their considerable efforts and time invested into learning English (Kim, 2005). Douglas (2000) also claimed that the TOEIC is "a good example of a well-constructed norm-referenced traditional multiple-choice test task, with no doubt high reliability, but extremely limited in the inferences it will allow about language knowledge" (p. 236). Thus, it seems that the TOEIC listening and reading scores are very limited in indicating candidates' real communicative business ability in real world situations. Instead, the test assesses test takers' English listening and reading comprehension ability in daily and business situations.

In order to compensate the limitations of the TOEIC listening and reading test, the TOEIC speaking and writing tests were developed in order to directly assess the ability to speak and write in English in a workplace setting (Powers, Kim, Yu, Weng, & Van Winkle, 2009). In this study, the TOEIC writing test is reviewed.

The TOEIC writing tests is independent from the other TOEIC tests (ETS, 2012). It consists of eight tasks and takes approximately one hour to complete (ETS, 2012). The task difficulty increases as the test taker progresses through the test (ETS, 2012). The following table 2 shows how the tasks in the TOEIC writing test are organized (ETS, 2012, p. 3).

Table 2. Composition of the TOEIC Writing Test

Question	Task	Evaluation Criteria
1-5	Write a sentence based on a picture	-Grammar -Relevance of the sentences to the picture
6-7	Respond to a written request	-Quality and variety of your sentences -Vocabulary -Organization
8	Write an opinion essay	-Whether the opinion is supported with reasons and/or examples -Grammar -Vocabulary -Organization

In terms of reporting the test score, the TOEIC writing tests' responses are scored based on a scale of 0 to 200 (ETS, TOEIC, 2015). Scoring is conducted based on the online scoring network, which is a scoring system that enable ETS to score anonymously (ETS, TOEIC, 2015). As a way of contributing to reliability, at least six different raters contribute to the final score of a candidate who takes speaking and writing tests (ETS, TOEIC, 2015).

While the TOEIC writing test is a performance-based writing test, it seems to be appropriate to assess the second language proficiency rather than "competency-based assessment (CBA), which emphasizes the competencies required by the learner to perform a particular job adequately" (Davies et al., 1999, p. 27). Thus, similar to the TOEIC listening and reading test, the TOEIC writing test does not assess business knowledge despite its business and commercial appearance. Although the TOEIC examinee handbook states a particular emphasis communication in business, commerce, and industry, it is unlikely that the listening, reading, and writing tasks engage the test takers in genuinely communicative behavior or genuinely specific purpose language use, and the test would be placed near the general end of the speciality continuum. Nevertheless, it seems that specific knowledge of constructing an English email and an argumentive essay is necessary to perform well in the writing task, as these types of written genre are asked to successfully complete the task.

## 2.2. The Oxford International Business English Certificate (OIBEC)

Another standardized test of business English is the Oxford International

Business English Certificate (OIBEC), designed by the University of Oxford Delegacy of Local Examinations. It is aimed at business people working in international commerce who intend to obtain a certificate of competence in English language skills for the purposes of promotion or changing employment (University of Oxford Delegacy of Local Examinations, 1990). The test is available in some 44 countries world-wide (Douglas, 2000), even though it is not offered in Korea. The test is given at two levels: First and executive level. The test takes about 125-145 minutes and tests all four skills (Douglas, 2000). The reading and writing task types from the OIBEC examination is summarized in Table 3 (O'Sullivan, 2006).

Table 3. Reading and Writing Task Types From the OIBEC

	First level	Executive level
Preparation	package	
Content	4 pages of written input: Range from report to letter to table and graphic	6/7 page of written input: Consists of a detailed contextualization, with excerpts from reports, letters, balance sheets, memos etc.
Reading and writing		
Time allotted	75 minutes	95 minutes
Marks awarded	100	100
Task 1	Reading based on preparation materials	Writing: Report completion (based on Prep. Materials): 2 pages allowed, 20 marks
Task 2	Reading: Inferencing	Writing: Guided report (based on Prep. Materials): 2 pages allowed, 15 marks
Task 3	Writing: Register	Reading: 5 items, based on additional fax input, 15 marks
Task 4	Writing/Reading integrated: Table completion/summary	Proofreading: 10 items in short memo text, 10 marks
Task 5	Guided writing: Memo, no word limit	Writing: Briefing paper completion, 2 paragraphs, 20 marks
Task 6	Letter writing-no word limit	Writing: Job application letter, 150-200 words, 20 marks

An interesting characteristic of the reading and writing section on the OIBEC was the inclusion of an extensive preparation package, which was

provided to every candidate three days before the day of the examination (O'Sullivan, 2006). The preparation package includes extensive information that involves narratives, tablets, letters, memos, and other printed input, and candidates may take it and any notes they might make about it into the examination room, along with a dictionary (Douglas, 2000). As demonstrated in Table 3, the reading and writing task types are based on a high degree of specificity. In terms of evaluation, test takers are told that they will be tested on ability to write with clarity, conciseness, and reasonable accuracy, their ability to set out a business letter, and their competency to solve problems (Douglas, 2000). Nevertheless, little information is given about criteria or procedures for scoring.

In comparison with the TOEIC, it appears that the reading and writing section on the OIBEC examination is a fairly well defined specific purpose test with a high degree of authenticity, particularly with interactional authenticity (Douglas, 2000). The input data is extensive and provides the examinees with complex business problems that form the basis of the test tasks and engage an appropriate discourse domain (Douglas, 2000). Thus, the OIBEC seems to contribute to testing business English in terms of moving along the specificity continuum to a situation where the test had been based on a high degree of specificity (O'Sullivan, 2006). However, as the test was developed for business people who already had a certain amount of professional experience and is aimed at promoting or changing employment (University of Oxford Delegacy of Local Examinations, 1990), the test seems un appropriate for potential employees who do not have professional experiences in business fields.

### 2.3. The Business Language Testing Service (BULATS)

More recently, the Business Language Testing Service (BULATS) has been newly produced by Cambridge ESOL (Cambridge English Language Assessment, 2013). The BULATS is developed to assess test takers' foreign language ability in the work place, and it offers a number of independent tests, such as The BULATS Standard test, the BULATS Computer test, the BULATS writing test, and the BULATS speaking test. Since the format of the BULATS computer test reflects that of the Standard test, the BULATS Standard test (Table

4) and the BULATS writing test (Table 5) are briefly reviewed in the following (Cambridge English Language Assessment, 2013).

Table 4. Composition of the BULATS Standard Test

Part	Format	Content	K
Listening section			
1	MCQ	Matching audio description to visuals or short phrases	10
2	SAF	Memo/form completion	12
3	Matching	Identifying speaker from list of topics/jobs etc.	10
4	MCQ	Listen for detail	18
Reading and language knowledge section			
<i>Part 1</i>			
1	MCQ	Understanding notices, messages, timetables, graphs, etc.	7
2	MCQ	Grammar and vocabulary. Gapped sentences tasks	6
3	MCQ	Reading long texts	6
4	Cloze	Grammar. Medium-length text	5
<i>Part 2</i>			
1	Matching	Statement to short texts	7
2	MC cloze	Reading medium-length text	5
3	Cloze	Grammar. Reading medium-length text	5
4	MCQ	Grammar and vocabulary. Gapped sentence tasks	6
5	MCQ	Reading long texts	6
6	SAF	Error recognition. Medium-length text.	7

*Note.* K= number of questions. MCQ= multiple choice questions. SAF=short answer format questions. MC cloze= multiple choice cloze questions.

The BULATS Standard test lasts 110 minutes, and its format contains various task types, such as multiple-choice, cloze, matching, and short answer questions (Cambridge English Language Assessment, 2013). The Reading and Language Knowledge section consists of two sections. Compared with Part 1, Part 2 is comprised of a higher-level task than Part 1 (Cambridge English Language Assessment, 2013). In terms of scoring, the BULATS provides six levels from Level 0 (score ranges between 10-19: beginner) to Level 5 (score ranges between 90-100: upper advanced) (Cambridge English Language Assessment, 2013). The levels are expressed based on Association of Language Testers in Europe (ALTE)

levels and the Council of Europe's Common European Framework of Reference for Languages (CEFR) (Cambridge English Language Assessment, 2013).

Besides the BULATS offers independent online writing subtests as demonstrated below in Table 5.

Table 5. Composition of the BULATS Writing Tests

Part	Title	Time	Focus
1	Short writing (50-60 words)	15 mins	Ability to write a short message, e-mail, fax or letter, covering given relevant points and using appropriate style and tone.
2	Extended writing (180-200 words)	30 mins	Ability to write a letter or report, using appropriate style and tone for the intended reader.

The writing test is composed of two tasks and lasts for 45 minutes. In the first part, candidates are asked to write a short text with a set of guidelines for writing a reply or follow-up letter (Cambridge English Language Assessment, 2013). Candidates are expected to write about all the points in the directions within about sixty words (O'Sullivan, 2006). The second task offers task takers a choice of either an extended letter or a report (Cambridge English Language Assessment, 2013).

With regard to the test report, the BULATS provides a candidate's overall test score on a scale of 0-100, levels based on ALTE as well as CEFR, and their score of each section of the test (e.g., Reading and Listening, and Writing) (Business Language Testing Service, 2013). In terms of scoring criteria in the writing section, performance on the tasks is assessed by two trained and accredited examiners working independently of one another (O'Sullivan, 2006). The criteria used are accuracy and appropriacy of grammar and vocabulary, organization of ideas, and achievement of purpose (Business Language Testing Service, 2013; O'Sullivan, 2006).

In comparison with the TOEIC listening and reading test, the format and content of the BULATS standard test are generally similar to the TOEIC listening and reading test, but the BULATS seems to have more variation of tasks and formats than the TOEIC listening and reading test. In addition, on the BULATS standard test, it is worth noting that grammar and vocabulary

knowledge is of great importance, containing 34 questions out of 60 total questions in the reading and language knowledge section. Regarding the degree of specificity, even though the listening part clearly focused on business-oriented contextualization, it seems that it was mainly concerned with social language. In comparison with the listening section, the reading and language knowledge section is considered less focused on business related text types.

In terms of the test appropriateness of use in Korea, the BULATS Standard test seems to be replaced by the TOEIC listening and reading test. The BULATS Standard test has been offered to measure business English proficiency in Korea. Nevertheless, it is interesting to note that Korean test takers are far less familiar with tests developed by Cambridge ESOL than those developed by the ETS.

On the contrary, the BULATS writing test seems inappropriate for Korean business contexts. In particular, it should be noted that the BULATS writing test generally covers topics and situations of general interests in offices, general business environments, and routines in ESL contexts, which are not always relevant to the Korean business context.

### 3. Current Issues in LSP Testing

#### 3.1. The Washback Effects of the LSP Tests

The ever-increasing use of the LSP testing around the globe has an effect on teaching paradigms and educational system as well as individual learning styles and future careers (Spolsky, 1997; Wall, 1998). In particular, a certain large-scale testing such as the TOEIC listening and reading test continues to be administered with more force than ever before by government and educational system throughout the world (Choi, 2008). Therefore, their impacts have resulted in many negative washback effects.

As many studies on washback effects of the traditional testing have predominantly demonstrated, the large-scale LSP testing narrowed the curriculum by forcing teachers to teach to the test (teach only the subject area) and to exclude non-testing subjects, such as writing, from their curriculum (Berry, 2008; Brimijoin, 2005; Gerwin & Visone, 2006; Renter et al., 2006; Shin, 2013; Sloan, 2005; Tsai &

Tsou, 2009; Van Hover, 2006; Wright & Choi, 2005). This kind of rigorous test-driven classroom environment decreased students' motivation in learning (Jung, 2008; Klinger & Luce-Kapler, 2007) and restricted their learning benefits (Brookhart & Bronowicz, 2003). Furthermore, the high-stakes LSP testing led teachers to have a fractured form of knowledge (Agee, 2004; Gerwin & Visone, 2006; Sloan, 2005; Smith, 2006; Van Hover, 2006). A fractured form of knowledge was shown in "the teaching of content in small, individuated, and isolated test-size pieces, as well as teaching in direct relation to the tests rather than in relation to other subject matter knowledge" (Au, 2007, p. 262).

Moreover, a lot of studies about student perspectives of LSP testing demonstrated problems related to the tests: too broad of a scope of test contents to prepare for (Teemant, 2010; Tsai & Tsou, 2009), culturally unfamiliar test topics (He & Shi, 2008), boring and dull processes of test preparation (Klinger & Luce-Kapler, 2007; Triplett & Barksdale, 2005; Tsai & Tsou, 2009), and most importantly, students' views on the test score (He & Shi, 2008; Teemant, 2010; Tsai & Tsou, 2009). That is, the students believed that the large-scaled LSP testing did not yield scores that reflected their true ability in English because they relied on their memorizations of writing samples rather than their own writing skills (He & Shi, 2008).

### **3.2. The Washback Effects of the TOEIC in Korea**

The negative washback effects in the LSP tests are even more true in the Korean society. In particular, as the high scores on the TOEIC listening and reading test have played a predominant role in determining the future success of current and potential employees, a number of Korean people generally invest enormous time and energy in achieving high scores on the test beginning as early as elementary education. Choi (2008) revealed that some Korean elementary school students (n=28 out of 100) have even taken the TOEIC listening and reading test (n=16) and the TOEFL (n=12), which is considered too difficult for their language proficiency and cognitive levels. These students took these tests because of the influence of their parents and the instructors of cram schools. From educational perspectives, however, such influence to take these LSP tests are unethical and unacceptable (Hamp-Lyons, 1997; Norton, 1997).

Misuse of the TOEIC listening and reading test is another important issue in Korea. Choi (2008) revealed that many Korean universities require students to achieve a certain range of LSP testing scores such as the TOEIC listening and reading test and TOEFL as graduation requirement (Choi, 2008). However, such a graduation requirement is criticized by many students and professors, since “these tests may not serve the purpose of graduation requirements adequately” (Choi, 2008, p. 57). Moreover, for graduate school admission, most Korean universities require candidates to submit the TOEFL or TOEIC listening and reading score (Choi, 2008). Some state examinations, such as the bar exam also require candidates to achieve a certain level of LSP tests (Choi, 2008). However, a number of English scholars in Korea claimed that the TOEIC listening and reading tests should not be used as an indicator of candidates’ overall English proficiency for the bar exam because the TOEIC listening and reading test only aims at measuring English listening and reading abilities in business related contexts (Choi, 1997, 2008; Lee, 2014).

In addition, several studies emphasized the problem of the TOEIC listening and reading test in the Korean business context. Park et al. (1998) revealed that Korean company executives questioned the scoring of the TOEIC listening and reading because it does not provide information on candidates’ authentic English abilities. Of 812 participants surveyed, 96.68 percent believed that “there is little correlation between the TOEIC listening and reading test score and English ability” (p. 184), and Park et al. (1998) further indicated that “the existing TOEIC listening and reading test has the problem of content validity and construct validity” (p. 184). Kim (2005) also reported significant difference between test takers’ TOEIC listening and reading test scores and actual English abilities. Kim (2005) stated that “a high TOEIC listening and reading score does not always guarantee job candidates’ fluency in English” (para. 5).

Overall, the negative washback effects of the LSP testing and those in the Korean testing contexts were discussed. In particular, considering that the Korean industry has grown quickly and has created a great demand for international communication in the field of business, it is highly problematic that the TOEIC listening and reading test is still misused as a prerequisite testing tool to measure test takers’ overall English ability. Thus, the LSP testing should be used in accordance with the purpose of the test.

## 4. Future Directions in LSP Testing

Even though large-scaled LSP testing has significantly influenced not only English teaching and learning, but also the social structure in general, it seems obvious that there is little initiative in terms of a critical perspective on the uses of large-scaled LSP testing, and social impacts on the testing. In particular, as pointed out in the previous section, it is clear that LSP testing is increasingly used for unintended purposes. For instance, in spite of being developed solely as listening and reading test of English for international communication, the TOEIC listening and reading test has been misused as an indispensable testing tool to assess candidates' overall English proficiency in the Korean society. This phenomenon is highly problematic and it would be even more problematic if no one provides any alternative solution to the issue of LSP testing and its negative impacts.

Unethical use of LSP testing is another important issue of study in the future. As previously pointed out, the TOEIC listening and reading test and the TOEFL test have been recently used for young students in Korea as a tool to motivate them to study English and enhance their English ability (Choi, 2008). Most young students were being made by adults to take LSP testing without any serious attempt to validate LSP testing for its purpose (Choi, 2008, O'Loughlin, 2008). The uses of both tests were thus considered by many language testing specialists to be unethical. Therefore, there is a clear need for inquiries designed for the consequence of the test from test takers' perspectives.

Furthermore, based on the review of LSP testing in this study, there is a crucial need for assessments designed locally for the different purposes. Park et al. (1998) pointed out that the employees (n=812) in major Korean companies clarified the need for developing an ESP test (mean=4.07) that is specified in the Korean context (Park et al., 1998). Lee's (2014) study moved one step forward in terms of the needs analysis in LSP testing. Lee (2014) conducted context specific need analysis intended for Korean workplace and analyzed frequently used business English writing in the field of an international sourcing, investment financing, and foreign banking. It is noteworthy that the study revealed a huge gap between business writing used in real workplace and writing tasks used in LSP testing. Furthermore, it is interesting to note that there is only few places to

learn business English writing systematically in Korean higher educational institutions even though business English writing is used in diverse business contexts. Nevertheless, need analysis and content analysis conducted in Lee (2014) study only includes limited amount of international business areas. Thus, it is worth consideration of conducting content analysis in more diverse business fields although it seems very challenging to examine all different types of business English writing used in Korean business fields.

Finally, since performance-based LSP tests, such as the TOEIC speaking and writing tests, have been recently introduced in Korea, major Korean companies increasingly requires these test scores to job candidates (Lee, 2012). However, there is little initiative regarding the large-scaled performance testing, such as analysis of test contents, test takers' perspective on the test, test consequences, socio-political and power dimensions of tests, and the relationship among the TOEIC listening, reading, speaking and writing skills in Korean academia. Thus, in the future, it would be worth taking consideration of conducting these areas of studies.

## References

- Agee, J. (2004). Negotiating a teaching identity: An African American teacher's struggle to teach in test-driven contexts. *Teachers College Record*, 106(4), 747-774.
- Au, W. (2007). High-stakes testing and curriculum control: A qualitative metasynthesis. *Educational Research*, 36(5), 258-267.
- Berry, R. (2008). *Assessment for learning*. Hong Kong, China: Hong Kong University Press.
- Brimijoin, K. (2005). Differentiation and high-stake testing: An oxymoron? *Theory Into Practice*, 44(3), 254-261.
- Brookhart, S. M., & Bronowicz, D. L. (2003). 'I don't like writing. It makes me fingers hurt': Students talk about their classroom assessments. *Assessment in Education*, 10(2), 221-242.
- Business Language Testing Service (2013, February 15). Results. Retrieved March 3, 2015, from <http://www.bults.org/computer->

based-tests/results

- Cambridge English Language Assessment (2013, February 16). Exams and qualifications. Retrieved March 3, 2015, from <http://www.cambridgeenglish.org/exams-and-qualifications/>
- Choi, I. C. (1997). Essential test method facets of a general English proficiency test and their validity as perceived by test-takers. *Language Research*, 33(4), 773-799.
- Choi, I. C. (2008). The impact of EFL testing on EFL education in Korea. *Language Testing*, 25(1), 39-62.
- Douglas, D. (2000). *Assessing languages for specific purposes*. Cambridge, MA: Cambridge University Press.
- ETS (2012). *TOEIC Examinee Handbook: Speaking & Writing*. Princeton, NJ: Educational Testing Service.
- Education Testing Service: TOEIC (2013, March 11). The TOEIC test-the global standard for assessing English proficiency for business. Retrieved March 1, 2015, from [http://www.ets.org/toEIC/succeed?WT.ac=toeichome\\_succeed\\_121127](http://www.ets.org/toEIC/succeed?WT.ac=toeichome_succeed_121127)
- Education Testing Service: TOEIC (2015, April 1). TOEIC scores. Retrieved March 1, 2015, from [http://www.ets.org/toEIC/speaking\\_writing/scores/](http://www.ets.org/toEIC/speaking_writing/scores/)
- Gerwin, D., & Visone, F. (2006). The freedom to teach: Contrasting history teaching in elective and state-tested courses. *Social Education*, 34(2), 259-282.
- Hamp-Lyons, L. (1997). Washback, impact and validity: Ethical concerns. *Language Testing*, 14(3), 295-303.
- He, L., & Shi, L. (2008). ESL students' perceptions and experiences of standardized English writing tests. *Assessing Writing*, 13(2), 130-149.
- Jung, H. (2008). The washback effects of English listening test in Korean CSAT on secondary school English learning and teaching. *The Linguistic Association of Korea Journal*, 16(3), 143-162.
- Jung, H. (2010). Washback effects of new TOEIC on Korean college students' English teaching and learning. *Mirea English Language and Literature Society*, 23(3), 193-207.
- Kim, H. K. (2005, May 10). Education Column: TOEIC. *Hankookilbo*. Retrieved

- March 3, 2015, from <http://weekly.hankooki.com/1page/nation/200505/wk2005052615132237070.html>
- Klinger, D. A., & Luce-Kapler, R. (2007). Walking in their shoes: Students' perceptions of large-scale high-stakes testing. *The Canadian Journal of Program Evaluation*, 22(3), 29-52.
- Lee, G. J. (2012, September 10). TOEIC speaking test. Retrieved March 2, 2015, from <http://eto.co.kr/news/outview.asp?Code=20110316101157277&ts=100615>
- Lee, K. J. (2014, June 24). Power elites. *Mediadaum*. Retrieved March 5, 2015, from <http://media.daum.net/politics/newsviewsnewsid=20140221031854881>
- Lee, Y. K. (2014). *Understanding the Korean culture of assessment and implementation of a performance assessment tool*. Unpublished doctoral dissertation, IUP, Indiana, PA.
- McNamara, T. (1996). *Measuring second language performance*. London, United Kingdom: Longman.
- Norton, B. (1997). Accountability in language assessment. In C. Clapham & D. Corson (Eds.), *Encyclopedia of language and education: Volume 7, language testing and assessment* (pp. 313-322). Norwell, MA: Kluwer Academic Publishers.
- O'Loughlin, K. (2008). Assessment at the workplace. In E. Shohamy & N. H. Hornberger (Eds.), *Language testing and assessment: Encyclopedia of language and education, volume 7* (pp. 69-80). New York, NY: Springer.
- O'Sullivan, B. (2006). *Issues in testing business English: The revision of the Cambridge business English certificates*. Cambridge, MA: Cambridge University Press.
- Park, N. S., Cho, J. H., Kim, C. B., Kim, E. D., & Choi, I. C. (1998). Study about practical English and communicative education. *English Education*, 53, 179-201.
- Powers, D., Kim, H. J., Yu, F., Weng, V., & VanWinkle, W. (2009). The TOEIC Speaking and Writing Tests: Relations to Test-Taker Perceptions of Proficiency in English. ETS Research Report.
- Renter, D. S., Scott, C., Kober, N., Chudowsky, N., Joftus, S., & Zabala, D. (2006). *From the capital to the classroom: Year 4 of the no child left behind act*. Washington, DC: Center on Education Policy.

- Shin, I. Y. (2013). The needs for EAP programs in the Korean tertiary institutions: Perceptions of Korean science students in the UK. *The Linguistic Association of Korea Journal*, 21(2), 139-162.
- Sloan, K. (2005). Playing to the logic of the Texas accountability system: How focusing on "ratings"-not children-undermines quality and equity. In A. Valenzuela (Ed.), *Leaving children behind: How "Texas-style" accountability fails Latino youth* (pp. 153-178). Albany, NY: State University of New York.
- Smith, A. M. (2006). Negotiating control and protecting the private: History teachers and the Virginia standards of learning. In S. G. Grant (Ed.), *Measuring history: Cases of state-level testing across the United States* (pp. 221-247). Greenwich, CT: Information Age.
- Spolsky, B. (1997). The ethics of gatekeeping tests: What have we learned in a hundred years? *Language Testing*, 14(3), 242-247.
- Teemant, A. (2010). ESL student perspectives on university classroom testing practices. *Journal of the Scholarship of Teaching and Learning*, 10(3), 89-105.
- Triplett, C. F., & Barksdale, M. A. (2005). Third through sixth graders' perceptions of high-stakes testing. *Journal of Literacy Research*, 37(2), 237-260.
- Tsai, Y., & Tsou, C. H. (2009). A standardized English language proficiency test as the graduation benchmark: Student perspectives in higher education. *Assessment in Education: Principles, Policy & Practice*, 16(3), 319-330.
- University of Oxford Delegacy of Local Examinations (1990). *OIBEC: Executive level syllabus and specimen material*. Oxford, England: University of Oxford.
- Van Hover, S. D. (2006). Teaching history in the old domination: The impact of Virginia's accountability reform on seven secondary beginning history teachers. In S. G. Grant (Ed.), *Measuring history: Cases of state-level testing across the United States* (pp. 195-219). Greenwich, CT: Information Age.
- Wall, D. (1998). Impact and washback in language testing. In C. M. Clapham & D. Corson (Eds.), *Language testing and assessment: Encyclopedia of language and education, volume 7* (pp. 291-302). Dordrecht: Kluwer Academic.
- Wright, W. E., & Choi, D. (2005). *Voices form the classroom: A statewide survey of experienced third-grade English language learner teachers on the impact of language and high-stakes testing policy in Arizona*. Tempe, AZ: Educational Policy Studies Laboratory, Arizona State University.

Yo, I., & Rie, K. (2011). Factor structure of the revised TOEIC test: A multiple-sample analysis. *Language Testing*, 29(1), 131-152.

**Youn-Kyoung Lee**

Department of English Language and Literature  
College of Humanities, Catholic University of Daegu  
Hayang-ro 13-13, Hayang-eup  
Gyeongbuk 712-702, Korea  
Phone: (053)850-3391  
Email: younpan1@cu.ac.kr

Received on May 1, 2015

Revised version received on September 2, 2015

Accepted on September 30, 2015