

Paper must be double-line spaced throughout. Garamond font, size 12 should be used throughout this includes all references, and tables. Left, right, top & bottom should have 1 inch margin.

An Analysis of Testing Practices in College Korean Language Classrooms¹

Title must be Bold and Centered.

Hye-Sook Wang
Brown University

Single space between title and author. Author names should be in bold. University should be in italics. No space between author and university. Single space between authors if more than 1

Abstract

The aim of this study is to examine current testing practices in Korean language colleges and universities. Two beginning and intermediate level listening Bachman & Palmer's 'task' analytical tool. The results show that first, the majority of tests includes listening, reading, and writing, while a few tests focus exclusively on grammar and vocabulary. Second, the relative proportion of the four skills varies notably from school to school and varies less between levels. Third, the sequence of parts/tasks is very similar in all tests regardless of program or level. Fourth, selected response and limited production are the most popular responses in grammar/vocabulary section of all the samples. There are virtually no extended response questions. Fifth, the length of input in reading and listening passages varies notably from program to program, although the results are hard to generalize for level differences. Implications of these findings are discussed.

Keywords: Teaching Korean as a Foreign Language; Assessment and testing; Korean language classrooms; Achievement test; Final examination

Keyword section should be single line space after Abstract section

¹ The original version of this paper was presented at the annual conference of the American Association of Teachers of Korean at Boston University in June 2014.

Footnotes should use Garamond font size 8

Introduction

Assessment is one of the most important aspects and crucial activities that any teachers engage in, be it through tests or other means. Teachers spend a lot of time both in developing the assessment (rubrics) as well as grading/evaluating. A number of fundamental questions need to be addressed: a) how do teachers make tests as clear and unambiguous as possible, and truly ‘useful’ and ‘meaningful’ to serve the intended purposes; b) how do teachers measure success in terms of the relationship between teaching (i.e., what they instruct) and testing (i.e., what they evaluate/assess), especially in classroom-based achievement tests; c) how can challenges in test development be effectively addressed and overcome from a practical perspective; and d) how do teachers assess students’ actual learning as well as their knowledge or skill of testing—two separate issues. The relative importance of tests in a given course may depend on various factors, such as the level and objectives of the course. But a general practice for many language teachers as described in their course syllabi appears to be that tests (e.g., unit/lesson tests, mid-term exams, final exams) are given relatively heavy weight, if not heaviest, compared to other course requirements such as attendance and participation, projects, homework, or quizzes, etc. This is especially true in the case of lower-level courses. Naturally, students seem to be gravely concerned about and sensitive to the impact of test scores on their overall course grade in addition to or irrespective of their actual learning outcomes. In this respect, fair and effective test administration on the part of the teachers as test developers—the entire process from design, development, to grading a test—becomes even more critical. Despite the significance of this topic in and of itself, little attention has been paid to this aspect in the field of Korean as

All headings should be in bold. There should be single line spacing between heading and paragraph after heading. All paragraphs after a heading should not be indented.

a Foreign Language (KFL henceforth), evidenced by a small volume of publications (see the literature review in the next section). While test development, administration, and grading are seen as a routine practice for teachers, perhaps one of the reasons any testing-related activities are taken for granted, it is time to critically reflect on current practices.

We analyze tests for a number of reasons and purposes. First, testing is one of the most essential and crucial components of teaching and learning processes in a classroom setting, and arguably the most laborious and challenging task for teachers. Other performance-based activities and tasks that learners produce (e.g., essays, projects, presentations, etc.), tests ride primarily on teachers' assessment knowledge and abilities. Second, tests function as a platform for useful discussion for teacher education for novice and experienced teachers alike. Being aware of good practices and making every effort to ensure tests are truly useful and meaningful for our students is an unavoidable responsibility for any teacher. Third, the effects of tests on instruction and curriculum, which is widely discussed in assessment literature as 'washback effect,' will guide teachers for any changes and/or adjustments needed.

Among various tests used in language classes, this paper specifically focuses on achievement tests. Achievement tests measure what a learner has learned or achieved during a course of instruction, typically in a form of mid-term or final exams as well as chapter/unit exams. However, the study focuses on 'final' achievement tests, not 'progress' achievement tests (Hughes, 2003). The final exam was chosen for analysis because it is a prototype summative test that concludes the work of a semester, as opposed to unit/lesson or chapter tests in which only focused materials (e.g., particular

Paragraphs (aside from the one immediately following a heading) should be indented and double line space between paragraphs

Table 1. Instructional Language

Level	NL	TL	Both	Mixed
1 st	6	0	0	0
2 nd	4	1	0	1

(Note: NL-learners' native language TL-learners' target language)

Use structured table instead of loose table structures. Loose structures don't hold well together during production editing. See example of structure table below

4.3.1.2. Structure

The structure includes number of parts/tasks, salience of parts/tasks, sequence of parts/tasks, relative importance of parts/tasks, and number of tasks/items per part. In regards to the number of parts or tasks in the test structure, the number ranges from 6 to 15 in the first-year test, and from 4 to 9 in the second-year test. The range is larger in the first-year than in the second-year. Two samples have 8 parts or tasks in the first year, and two samples have 4 parts or tasks in the second year, while the number of parts or tasks in all other samples varies. A general trend seems to be that there are considerably more parts in first-year than in second-year tests.

With respect to the salience of parts/tasks, different parts of the test are clearly distinguished from one another in almost all the tests in both the first-year samples and the second-year samples. Only one test in each level did not show any clear distinction. When the sequence of parts or tasks is analyzed, it is revealed that the test begins with listening and ends with writing in the majority of the samples. In the case of 1st year, three tests (50%) followed the order of listening-> vocabulary, expression, grammar -> reading -> writing. The other two followed a slightly different order, although keeping the pattern of beginning with listening and ending with writing. One test did not have listening or reading components. It was comprised of vocabulary, grammar, and expression (e.g. more than a single word). The second-year tests also followed a very similar pattern as the 1st year tests as the following tables display.

Table 2. Sequence of Parts/Tasks (1st Year)

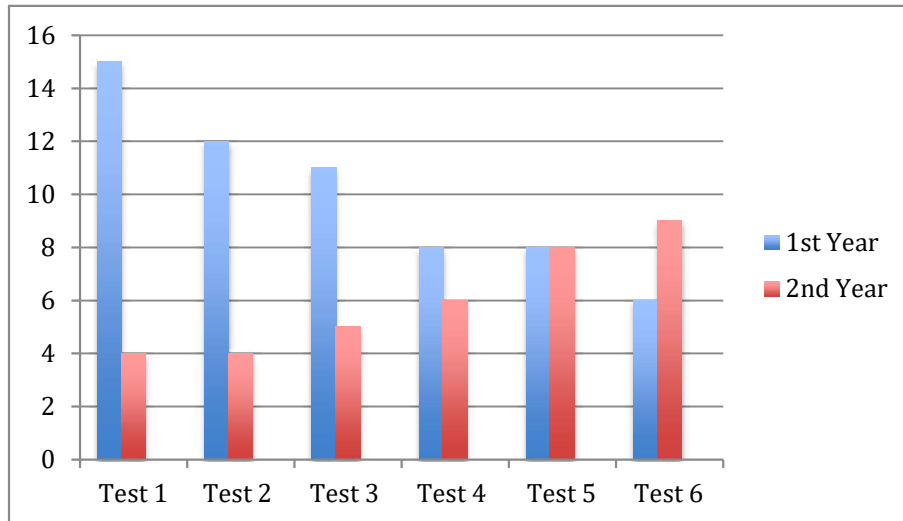
	L>VEG>R> W	L>VEG>R>G> W USE	L>R>VEG>L> W	VEG
	50%	17%	17%	17%

Example of structured table

(Note: total percentage is over 100 because of roundup.

L-Listening VEG-Vocabulary, Expression, Grammar R-Reading W-Writing)

Figure 1. Number of Parts/Tasks in Test Structure



Graphs should be inserted as pictures instead of real data driven graphs. Real graphs do not hold well together during production editing

1st Yr.: 1 test (15), 1 test (12), 1 test (11), 2 tests (8), 1 test (6) – ranges from 6 to 15

2nd Yr.: 2 test (4), 1 test (5), 1 test (6), 1 test (8), 1 test (9) – ranges from 4 to 9

References

- Allison, D. (1999). *Language testing and evaluation: An introductory course*. Singapore: Singapore University Press.
- American Association of Teachers of Korean. Retrieved on January 2020.
<http://www.aatk.org/www/html/textbooks.html>
- Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice: designing and developing useful language tests*. Oxford: Oxford University Press.
- Barrette, C. (2004). *Assessing language achievement*. *Language Annals*, 37(1), 58-70.
- Brown, H. D., & Abeywickrama, P. (2010). *Language assessment: Principles and classroom practices*. White Plains, NY: Pearson Education.
- Brown, J. & Hudson, T. (1998). The alternatives in language assessment. *TESOL Quarterly*, 32(4), 653-675.
- Chang, E. (2011). Korean language education assessment studies (Hankuke Gyoyuk Pyungka Yeongu). *Bilingual Research*, 47, 351-382.
- Cho, S. (2004). Oral Proficiency Interview: Pros and cons. *Korean Language in America*, 9, 144-151.
- Choi, E-G. (2006). The practice and task of the Korean language testing. *Journal of Korean Language Education*, 17 (2), 289-319.

References should use hanging indentation formatting.