Developing a Test for Assessing Listening through Text Patterns and Graphic Organization

Alan Orr and Kyujin Lee

Northern Arizona University

Abstract

In this project report, the development, administration, and results of a listening test that was created for English 638: Assessment for Second Language Skills is described. The test was created in order to answer two questions: The test developers wanted to know whether a valid, reliable, and practical test for assessing text patterns in listening passages could be developed. The test developers also wanted to know whether the results could provide formative assessment for PIE Listening and Speaking and Content-based Instruction classes. The results from the pilot administration of the test indicated that a somewhat valid, reliable, and practical test could be developed. Additionally, the test provided enough information to serve as formative assessment; the test developers recommended that the coordinators of Level 3 Listening and Speaking and CBI classes at the PIE consider explicitly teaching students to identify text patterns in listening passages and to arrange content from listening passages in partially completed graphic organizers.

Keywords: listening assessment, text patterns, graphic organizers

Developing a Test for Assessing Listening through Text Patterns and Graphic Organization

Background

The test developers were required to create a test for English 638: Assessment for Second Language Skills. They created a listening test focused on text patterns. It assessed what Bachman and Palmer (2010) call grammatical and textual knowledge from both a top-down and a bottom-up view of listening processing.

Research Questions

The test developers sought to answer two questions: First, they wanted to know whether a test focused on text patterns and arranging key ideas could be a valid, reliable, and practical way to assess listening comprehension. Second, the researches wanted to know whether the results would indicate that Level 3 students at the PIE needed more explicit instruction about text patterns to improve their listening comprehension.

Methods

In this methods section, the participants, the design of the test, the scoring procedures, and the test administration is described.

Participants

The participants who took the test were 14 Level 3 students enrolled Kyujin Lee's CBI class at the PIE.

Design of Test

The test included two parts, each pertaining to the same 12 listening passages. Part 1 measured test takers' ability to identify text patterns with matching items, and Part 2 measured their ability to understand relationships of key ideas through the use of partially completed

graphic organizers.

Scoring Procedures

The items for Parts 1 and 2 were dichotomously scored, and the scores were provided in the following ways: as a total score, a score for each part, and a score for each text pattern.

Administration

The pilot administration occurred on November 14 and 19, 2013.

Results

In this section, the statistics from the pilot administration of the TDPIRLP are presented.

Statistics

In Appendix A, the descriptive statistics including item difficulty (P) and item discrimination (D) values for all test items are presented. Additionally, the descriptive statistics including the number of test takers (N), the total amount of items per domain (K), the minimum (Min) and maximum scores (Max), the mean, the standard deviation (SD), the item reliability using Cronbach's Alpha (r), the coefficient of agreement (po) using Subkoviak's Approximate Values of the Agreement Coefficient (1988, p. 49), and the standard error of measurement (SEM) are presented.

Relevance to PIE and Second Language Learning

The first research question focused on whether the test was valid, reliable, and practical. Although there were challenges to the tests' validity especially during its administration, the test was somewhat valid in that it served its purpose of providing formative information about the test takers in Kyujin Lee's class and their ability to identify text patterns and understand relationships of key ideas in listening passages. Both Kyujin Lee and Debra Goldenstein noted that the test ranked the test takers in approximately the same manner as their typical classroom

performance. As for reliability, the test overall had a reliability of .84, which is within Miller, Linn, and Gronlund's (2009) range for teacher-created classroom tests. However, the reliability of certain parts of the test was lower. The test was practical in that that the test developers created short listening passages; creating longer passages may be more time consuming. If this test were to serve as a model for an Achievement Test or a Skills Assessment at the PIE, the test takers recommend that only one or two graphic organizers be used to measure test takers' ability to understand relationships of key ideas because graphic organizers with appropriate distractors are somewhat time consuming to create.

The second research question asked whether the test could provide formative information for listening instruction in Listening and Speaking and Content-based Instruction classes at the PIE. Based on the results, the test takers decided that intermediate students at the PIE have some cursory knowledge of text patterns in listening passages. However, students could benefit from further explicit instruction about text patterns and the use of graphic organizers to arrange content from listening passages. Further instruction would be likely to yield more masters of the objectives in the curriculum related to comprehending listening passages and organizing content from listening passages.

For more information about the test, contact Kyujin Lee (kl523@nau.edu) or Alan Orr (amo84@nau.edu) to request a copy of the full project report, which includes a copy of the test.

References

Bachman, L., & Palmer, A. (2010). *Language assessment in practice*. Oxford, UK: Oxford University Press.

Miller, D. M., Linn, R. L., & Gronlund, N. E. (2009). *Measurement and assessment in teaching* (10th ed.). Upper Saddle River, NJ: Pearson Education.

Appendix A
Item Statistics and Descriptive Statistics

Item Statistics

Part 1: Identifying Patterns	Part, Passage, and Item Number	P	D	Part 2: Understand Relationships of Key Ideas	Part, Passage, and Item Number	P	D
Cause-Effect	1.1.1	.71	.07	Cause-Effect	2.1.1	.79	.57
	1.5.5	.64	.36		2.1.2	.79	.26
	1.12.12	.57	-.14		2.1.3	.71	.82
					2.5.1	.86	.11
					2.5.2	.93	-.07
					2.5.3	.86	.34
					2.12.1	.14	-.03
					2.12.2	.79	.38
					2.12.3	.29	.05
Comparison-Contrast	1.4.4	.71	-.22	Comparison-Contrast	2.4.1	.57	.55
	1.6.6	.57	.46		2.4.2	.50	.52
	1.10.10	.71	.23		2.4.3	.50	.10
					2.6.1	.86	.68
					2.6.2	.50	.35
					2.6.3	.64	.32
					2.10.1	.43	-.06
					2.10.2	.71	.23
					2.10.3	.71	.20
Problem-Solution	1.2.2	.64	.39	Problem-Solution	2.2.1	.79	.75
	1.8.8	.64	.28		2.2.2	.79	.57
	1.11.11	.64	-.37		2.2.3	.86	.49
					2.8.1	.93	.69
					2.8.2	.86	.46
					2.8.3	.64	.48
					2.11.1	.50	-.12
					2.11.2	.50	.37
					2.11.3	.71	.60
Sequence	1.3.3	.57	.53	Sequence	2.3.1	.86	.52
	1.7.7	.14	.11		2.3.2	.86	.37
	1.9.9	.64	.73		2.3.3	.93	.69
					2.7.1	.43	.28
					2.7.2	.50	-.04
					2.7.3	.79	-.21
					2.9.1	.86	.14
					2.9.2	.79	.16
					2.9.3	.86	.71

Descriptive Statistics

Domain	K	Min	Max	Mean	SD	r	po	SEM
Total	48	14	40	32.21	7.36	.84	.80	2.94
Patterns	12	4	10	7.21	2.39	.55	.74	1.60
Relationships of Key Ideas	36	9	32	25.00	6.13	.84	.80	2.44
Cause-Effect	12	5	11	8.07	1.82	.37	.64	1.44
Comparison-Contrast	12	4	12	7.43	2.62	.64	.73	1.58
Problem-Solution	12	1	10	8.50	2.38	.62	.70	1.47
Sequence	12	4	11	8.21	2.16	.59	.71	1.38

Note. N=14