TOEIC Speaking Test: A Correlational Study and Test Takers’ Reactions

Masaya Kanzaki
Kanda University of International Studies

Reference Data:

The TOEIC Speaking test is a computer-based examination developed by Educational Testing Service in 2006. This relatively new test was examined in this study with two objectives. The first was to explore correlations between scores of the TOEIC Speaking test and the TOEIC Listening and Reading test. The second was to investigate test takers’ reactions to the TOEIC Speaking test in comparison to the TOEIC Listening and Reading test. The study was conducted over 4 years, during which data were collected from 313 university students. The speaking scores moderately correlated with the scores for listening, reading, and listening and reading combined. As for the participants’ reactions to the TOEIC Speaking test, they were generally positive.

Since the test is relatively new and has not been subject to much research, revealing what kind of test this is through a study was deemed important.

This study was done in two parts. First, correlations between the scores of the TOEIC Speaking test and the TOEIC Listening and Reading test were investigated. Since the TOEIC Listening and Reading test is easier and less expensive to administer, it is widely used for measuring English language proficiency. However, how closely the scores of this test reflect speaking ability remains unclear. Two major ETS-commissioned correlational studies have examined the relationships between these scores. One is Liao, Qu, and Morgan (2010), which reported a correlation of .66 between the speaking and listening scores and a correlation of .57 between the speaking and reading scores (N=12,099). The other is Liu and Costanzo (2013), which reported a correlation of .63 between the speaking and listening scores and a correlation of .54 between the speaking and reading scores (N=4,935). The current study investigated whether similar correlations can be found when Japanese university students take the tests.

The second part of this study examined the participants’ reactions to the TOEIC Speaking test. Bradshaw (1990) and Coniam (1999) emphasized the importance of considering test takers’ reactions when evaluating a test, arguing that useful insights can be gained by studying their feelings and opinions about the test. They called this aspect of investigation “consumer validity” and encouraged more research on the subject, as it had received very little attention in the language testing literature. Similarly, Shohamy (2001) pointed out that listening to test takers’ reactions can be a great source of information. Also, she warned that reactions to a test can be negative when the test takers have low trust in the test. The current study investigated test takers’ reactions to gain an insight into the TOEIC Speaking test and to identify any negative reactions.

To sum up, this study was conducted to answer the following research questions:

RQ1. To what extent does the TOEIC Speaking test score correlate with the listening score, reading score, and total score of the TOEIC Listening and Reading test?
The objectives of the study as well as the related procedures and requirements were explained to the participants before they signed a consent form. They understood that the TOEIC Speaking test and questionnaire would not affect their TOEIC course grades in any way but that the results of the TOEIC Listening and Reading test would affect their grades.

**Materials & Procedures**

The TOEIC Speaking test, the TOEIC Listening and Reading test, and a questionnaire were used in this study. The TOEIC tests were administered in the Institutional Program (IP), with which an institution sets the exam date, time, and location. They were given over 2 consecutive days; the TOEIC Listening and Reading test on the first day and the TOEIC Speaking test on the second day. All the participants took the TOEIC Listening and Reading test at the same time, but due to a lack of computers, the TOEIC Speaking test was taken at different times, in the morning and afternoon. The results of the TOEIC tests were provided by the Institute for International Business Communication, the administrator of the TOEIC in Japan. The scores were then analyzed using Microsoft Excel 2013 and IBM SPSS Statistics 23. The questionnaire was given to the participants on a paper form after they completed the TOEIC Speaking test on the second day. The survey data were then entered into a Microsoft Excel spreadsheet for analysis.

**TOEIC Tests**

The TOEIC Speaking test is a computer-based examination requiring test takers to sit in front of a computer screen while wearing a headset with a microphone. Instructions are provided on the screen and through the headset. Test takers speak into the microphone, and their speeches are recorded and sent to certified raters for evaluation. There are 11 questions in the test, and scores are given in the range of 0 to 200.

The TOEIC Listening and Reading test is composed of two subtests, the listening test and the reading test. Each subtest consists of 100 multiple-choice questions, and raw scores of between 0 and 100 are converted to scaled scores of between 5 and 495. The listening test has four parts, whereas the reading test has three. Details of the TOEIC Speaking test and TOEIC Listening and Reading test are shown in Appendix A.

As the ETS does not make item-by-item results of the TOEIC available, reliability indices for the three TOEIC tests used in this study could not be calculated. However, the ETS (Educational Testing Service, 2010) reported that the reliability of the TOEIC Speaking test was .80 based on the data from tests administered in the Public Testing
Program from January 2008 to December 2009. Also, the ETS (Educational Testing Service, 2013) reported that the reliability of the TOEIC Listening and Reading section scores across all forms of their norming samples was approximately .90.

TOEIC Speaking Test Questionnaire
A questionnaire was developed by the author of the study to elicit the participants’ reactions to the TOEIC Speaking test. Created in 2013 and modified in 2014 and 2015, the questionnaire was administered as three different versions over 4 years. Questionnaire items used in this study are shown in Appendices B and C. The questions were in a multiple-choice format, but spaces were provided for comments.

Results and Discussion
Descriptive Statistics for the TOEIC Scores
Table 1 shows the descriptive statistics for the scores of the TOEIC Speaking test and TOEIC Listening and Reading test. The participants were better at listening than reading; the average listening score was 100.83 points higher than the average reading score.

Table 1. Descriptive Statistics for the TOEIC Tests (N = 313)

<table>
<thead>
<tr>
<th>Test</th>
<th>Possible scores</th>
<th>Mean</th>
<th>SD</th>
<th>Min</th>
<th>Max</th>
<th>Skewness</th>
<th>Kurtosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>S</td>
<td>0–200</td>
<td>116.45</td>
<td>21.57</td>
<td>40</td>
<td>180</td>
<td>−0.28</td>
<td>0.62</td>
</tr>
<tr>
<td>L</td>
<td>5–495</td>
<td>372.83</td>
<td>57.81</td>
<td>170</td>
<td>495</td>
<td>−0.49</td>
<td>0.63</td>
</tr>
<tr>
<td>R</td>
<td>5–495</td>
<td>272.00</td>
<td>72.00</td>
<td>100</td>
<td>475</td>
<td>0.11</td>
<td>−0.61</td>
</tr>
<tr>
<td>LR</td>
<td>10–990</td>
<td>644.82</td>
<td>119.63</td>
<td>310</td>
<td>945</td>
<td>−0.08</td>
<td>−0.16</td>
</tr>
</tbody>
</table>

Note. S = TOEIC Speaking test; L = TOEIC Listening test; R = TOEIC Reading test; LR = L and R combined; Min = lowest score; Max = highest score.

Correlations
Table 2 shows correlations between the scores of the TOEIC Speaking test and the scores of the listening test, reading test, and listening and reading test combined. The speaking scores significantly correlated with the listening scores, reading scores, and the listening and reading scores combined (all ps < .001).

Table 2. Correlations Between Test Scores (N = 313)

<table>
<thead>
<tr>
<th></th>
<th>L</th>
<th>R</th>
<th>LR</th>
</tr>
</thead>
<tbody>
<tr>
<td>S</td>
<td>.48</td>
<td>.48</td>
<td>.52</td>
</tr>
</tbody>
</table>

Note. S = TOEIC Speaking test, L = TOEIC Listening test, R = TOEIC Reading test, LR = L and R combined. All ps < .001 (two-tailed)

Figure 1 is a scatterplot showing the relationship between the scores of the TOEIC Speaking test and TOEIC Listening and Reading test.
Kanzaki: TOEIC Speaking Test: A Correlational Study and Test Takers’ Reactions

Questionnaire
The original questionnaire was modified twice, and so three different versions were used over 4 years.

In Question 1, participants were asked how they had found the TOEIC Speaking test as compared to the TOEIC Listening and Reading test (see Table 3 for responses). Multiple answers were allowed for this question. Of the 313 respondents, 177 (57%) indicated that the TOEIC Speaking test was difficult and 109 (35%) indicated that it was demanding, which suggests that the TOEIC Speaking test might have been too difficult for them. However, 143 (46%) indicated that the TOEIC Speaking test was fun.

Table 3. Responses to Question 1 (N = 313; Multiple Answers Allowed)

<table>
<thead>
<tr>
<th>Easy</th>
<th>Difficult</th>
<th>Undemanding</th>
<th>Demanding</th>
<th>Fun</th>
<th>Boring</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>177</td>
<td>13</td>
<td>109</td>
<td>143</td>
<td>1</td>
</tr>
</tbody>
</table>

Note. Question 1: How did you find the TOEIC Speaking test as compared to the TOEIC Listening and Reading test (the regular TOEIC)?

In Question 2, participants were asked which of the two TOEIC tests they liked better (see Table 4 for responses). Among the 147 respondents who indicated a preference for the TOEIC Speaking test, 63 chose (43%) difficult and 100 (68%) chose fun for Question 1, whereas among the 86 respondents who indicated a preference for the TOEIC Listening and Reading test, 65 (76%) chose difficult and 10 (12%) chose fun.

Table 4. Responses to Question 2 (N = 313)

<table>
<thead>
<tr>
<th>TOEIC Speaking test</th>
<th>TOEIC Listening and Reading test</th>
<th>I like both equally</th>
<th>I don’t like either</th>
</tr>
</thead>
<tbody>
<tr>
<td>147</td>
<td>86</td>
<td>38</td>
<td>42</td>
</tr>
</tbody>
</table>

Note. Question 2: Which do you like better, the TOEIC Speaking test or the TOEIC Listening and Reading test (the regular TOEIC)?

In Question 3, participants were asked whether they were at a loss what to do during the TOEIC Speaking test due to being unfamiliar with its format (see Table 6 for responses). The average speaking scores of those who chose the three options (many times, a few times, not at all) were 108.85, 120.38, and 125.56, respectively. It seems that being at a loss what to do during the test negatively affected their scores.

Table 6. Responses to Question 3 (N = 313)

<table>
<thead>
<tr>
<th>Many times</th>
<th>A few times</th>
<th>Not at all</th>
</tr>
</thead>
<tbody>
<tr>
<td>122</td>
<td>156</td>
<td>36</td>
</tr>
</tbody>
</table>

Note. Question 3: Were you at a loss what to do during the TOEIC Speaking test due to being unfamiliar with its format?

In Question 4, participants were asked whether they wanted to take the TOEIC Speaking test again (see Table 7 for responses). About 85% of them indicated that they wanted to take it again, which might be a result of positive feelings towards the test.

Table 7. Responses to Question 4 (N = 313)

<table>
<thead>
<tr>
<th>Very much so</th>
<th>Yes, I do</th>
<th>I’m not sure</th>
<th>Not very much</th>
<th>Not at all</th>
</tr>
</thead>
<tbody>
<tr>
<td>118</td>
<td>149</td>
<td>13</td>
<td>26</td>
<td>7</td>
</tr>
</tbody>
</table>

Note. Question 4: Do you want to take the TOEIC Speaking test again? The choice I’m not sure was used only in the 2013 version of the questionnaire.

Table 5. Average Scores by Test Preference (n = 233)

<table>
<thead>
<tr>
<th>Preference</th>
<th>S</th>
<th>L</th>
<th>R</th>
<th>LR</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOEIC Speaking test (n = 147)</td>
<td>117.14</td>
<td>370.10</td>
<td>261.90</td>
<td>632.01</td>
</tr>
<tr>
<td>Listening and Reading test (n = 86)</td>
<td>113.02</td>
<td>376.22</td>
<td>285.29</td>
<td>661.51</td>
</tr>
</tbody>
</table>

Note. S = TOEIC Speaking test, L = TOEIC Listening test, R = TOEIC Reading test, LR = L and R combined.

Table 5 shows the average TOEIC scores of those who indicated a preference for one of the two TOEIC tests. The average speaking score was higher among those who indicated a preference for the TOEIC Speaking test, but the average listening and reading scores were higher among those who indicated a preference for the TOEIC Listening and Reading test.
Kanzaki: TOEIC Speaking Test: A Correlational Study and Test Takers’ Reactions

In Question 5, participants were asked whether they thought their scores would improve if they took the TOEIC Speaking test again (see Table 8 for responses). The question was intended to gauge the appropriateness of the test’s difficulty level on the assumption that if the test takers’ abilities were far below the level needed to perform the tasks in the test adequately, they would not feel they could perform better the next time. On the other hand, if their abilities were slightly below that level, they would feel they could improve their scores the next time. Most (254) respondents (81%) thought that their scores would improve the next time.

<table>
<thead>
<tr>
<th>Very much so</th>
<th>Yes, I do</th>
<th>I’m not sure</th>
<th>Not very much</th>
<th>Not at all</th>
</tr>
</thead>
<tbody>
<tr>
<td>56</td>
<td>198</td>
<td>20</td>
<td>37</td>
<td>2</td>
</tr>
</tbody>
</table>

Note. Question 5: Do you think your score will improve if you take the TOEIC Speaking test again? The choice I’m not sure was used only in the 2013 version of the questionnaire.

In Question 6, participants were asked how well they had been able to demonstrate their speaking ability in the test (see Table 9 for responses). More than 90% of the respondents chose either not very well or not at all. This was not only the result of their unfamiliarity with the test format; it was also due to the fact that they were not used to talking to a computer, according to some of their comments on the questionnaire. Some of the participants also mentioned that they had been distracted by the voices of fellow test takers sitting nearby.

<table>
<thead>
<tr>
<th>Perfectly</th>
<th>Quite well</th>
<th>Not very well</th>
<th>Not well at all</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>9</td>
<td>107</td>
<td>17</td>
</tr>
</tbody>
</table>

Note. Question 6: How well were you able to demonstrate your speaking ability in the test? This question was given to only 135 participants in 2015 and 2016.

In Question 7, participants were asked whether speaking to a computer was more difficult than speaking to a person (see Table 10 for responses). Out of 135 respondents, 97 (72%) indicated that speaking to a computer is more difficult.

<table>
<thead>
<tr>
<th>Very much so</th>
<th>Yes, I do</th>
<th>Not really</th>
<th>No, not at all</th>
</tr>
</thead>
<tbody>
<tr>
<td>38</td>
<td>59</td>
<td>36</td>
<td>2</td>
</tr>
</tbody>
</table>

Note. Question 7: Compared to speaking directly to a person, do you feel that words do not come as naturally as when speaking to a computer? This question was given to only 135 participants in 2015 and 2016.

In Question 8, participants were asked whether they thought the TOEIC Speaking test is good for measuring speaking ability (see Table 11 for responses). Out of 299 respondents, 188 (82%) indicated that they trusted this aspect of the test.

<table>
<thead>
<tr>
<th>Very much so</th>
<th>Yes, I do</th>
<th>Not really</th>
<th>No, not at all</th>
</tr>
</thead>
<tbody>
<tr>
<td>44</td>
<td>144</td>
<td>41</td>
<td>0</td>
</tr>
</tbody>
</table>

Note. Question 8: Do you think the TOEIC Speaking test is good for measuring speaking ability? This question was not included in the questionnaire used in 2013.

In Question 9, participants were asked for their general opinions about taking tests such as the TOEIC Speaking test and TOEIC Listening and Reading test (see Table 12 for responses), and multiple answers were allowed. The majority of the participants appeared to recognize the value of taking tests, as 110 indicated that such tests are useful for assessing English level and 81 indicated that tests motivate them to study English.

Answer | Number
---|---
They are useful for assessing my English level. | 110
They motivate me to study English. | 81
They make me less motivated. | 2
They are meaningless. | 2
They are challenging and therefore worth trying. | 34

Note. Question 9: What do you think of taking tests such as the TOEIC Speaking test and TOEIC Listening and Reading test? This question was given to only 135 participants in 2015 and 2016.
Conclusion

The first part of this study examined correlations among the TOEIC scores. The speaking scores significantly correlated with the listening scores, $r = .48$, and the reading scores, $r = .48$ (all $p < .001$). The correlations are lower than those reported in Liao et al. (2010) and Liu and Costanzo (2013), which could be due to smaller sample size. Another difference in this study is the lower correlation between the speaking and listening scores, when compared to the previous studies. The correlations between the speaking and listening scores and between the speaking and reading scores were the same, whereas the former was higher than the latter in both Liao et al. (2010) and Liu and Costanzo (2013). This could be due to the greater skewness of the listening scores (skewness = −0.49) than that of the reading scores (skewness = 0.11). Although the correlations found in this study are statistically significant, they are not high enough to suggest a strong relationship between the speaking scores and the listening and reading scores. Therefore, estimating speaking ability using TOEIC Listening and Reading test results is not practical.

The second part of the study examined the participants’ reactions to the TOEIC Speaking test. The reactions were generally positive: 143 out of 313 participants (46%) indicated that the TOEIC Speaking test was fun; 147 (47%) indicated that they liked the TOEIC Speaking test better than the TOEIC Listening and Reading test; 267 (85%) indicated that they wanted to take the TOEIC Speaking test again; 254 (81%) thought that their scores would improve if they took it again; and 188 out of 229 (82%) thought the TOEIC Speaking test is good for measuring speaking ability. However, some of the participants’ responses to the questionnaire can be regarded as negative: 177 out of 313 participants (57%) found the TOEIC Speaking test difficult; 86 (27%) indicated that they liked the TOEIC Speaking test better than the TOEIC Listening and Reading test better; 278 (89%) were at a loss what to do during the test due to being unfamiliar with its format; 124 out of 133 (93%) indicated that they were unable to demonstrate their speaking ability well in the test; and 98 out of 135 (72%) indicated that speaking to a computer is more difficult than speaking to a person. Although some negative reactions were detected, overall the TOEIC Speaking test seems to serve as a motivating factor for many learners.

One limitation of this study is that the participants were language majors, and so their speaking ability was higher than the national average. It remains to be seen whether similar results can be obtained with the inclusion of other types of learners. In addition, participation in this study was voluntary; therefore, those who generally do not like taking tests did not participate. Thus, the results of this study cannot be generalized to the general population.

Acknowledgment

This study was supported by JSPS KAKENHI Grant Number 25370727.

Bio Data

Masaya Kanzaki teaches at Kanda University of International Studies. His research interests include language testing, vocabulary acquisition, and corpus linguistics. <kanzaki-m@kanda.kuis.ac.jp>

References


Appendix A

Details of the TOEIC Speaking Test and TOEIC Listening and Reading Test

Details of the TOEIC Speaking Test

<table>
<thead>
<tr>
<th>Question no.</th>
<th>Task</th>
</tr>
</thead>
<tbody>
<tr>
<td>1, 2</td>
<td>Read aloud the text that appears on the screen.</td>
</tr>
<tr>
<td>3</td>
<td>Describe the picture on the screen.</td>
</tr>
<tr>
<td>4-6</td>
<td>Answer three questions about a single topic as though you are participating in a telephone interview.</td>
</tr>
<tr>
<td>7-9</td>
<td>Read the information on the screen and answer three questions about it as though you are responding to a telephone inquiry.</td>
</tr>
<tr>
<td>10</td>
<td>Listen to a recorded message about a problem and propose a solution for it.</td>
</tr>
<tr>
<td>11</td>
<td>Express an opinion about a specific topic.</td>
</tr>
</tbody>
</table>

Details of the Listening Subtest of the TOEIC Listening and Reading Test

<table>
<thead>
<tr>
<th>Part</th>
<th>Task</th>
<th>No. of Qs</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>For each question with a photo, listen to four sentences and choose the one that best describes the image.</td>
<td>10</td>
</tr>
<tr>
<td>2</td>
<td>Listen to a question or statement followed by three responses and choose the most appropriate response.</td>
<td>30</td>
</tr>
<tr>
<td>3</td>
<td>Listen to a conversation and answer comprehension questions.</td>
<td>30</td>
</tr>
<tr>
<td>4</td>
<td>Listen to a short talk and answer comprehension questions.</td>
<td>30</td>
</tr>
</tbody>
</table>

Appendix B

TOEIC Speaking Test Questionnaire (Japanese)

Q1. TOEICスピーキングテストをTOEICリスニング・リーディングテスト（普通のTOEIC）と較べて、どう感じましたか？（複数回答可）
   1) 易かった    2) 難しかった    3) 楽だった
   4) 大変だった    5) 楽しかった    6) つまらなかった
Q2. TOEICスピーキングテストとTOEICリスニング・リーディングテスト（普通のTOEIC）とどちらが好きですか？
   1) TOEICスピーキングテスト    2) TOEICリスニング・リーディングテスト
   3) 両方好き     4) 両方嫌い
Q3. TOEICスピーキングテストの問題形式に慣れていないため、試験中、戸惑うことはありましたか？
   1) かなりあった    2) 少しあった    3) まったくなかった
Q4. TOEICスピーキングテストをまた受けたいですか？
   1) 強くそう思う    2) そう思う    3) どちらとも言えない
   4) あまりそう思わない    5) まったくそう思わない
Q5. TOEICスピーキングテストを次に受けたら、スコアは上がると思いますか？
   1) 強くそう思う    2) そう思う    3) どちらとも言えない
   4) あまりそう思わない    5) まったくそう思わない
Q6. 今回のTOEICスピーキングテストで自分のスピーキング力をどのくらい出せたと思いますか？
   1) 完全に出せた    2) かなり出せた
   3) あまり出せなかった    4) まったく出せなかった
Kanzaki: TOEIC Speaking Test: A Correlational Study and Test Takers’ Reactions

Q7. 直接人に向かって話す場合と較べて、コンピュータに向かって話すTOEICスピーキングテストでは言葉が上手く出てこないと感じましたか?
1) 強くそう感じた 2) そう感じた 3) あまりそう感じなかった 4) まったくそう感じなかった
Q8. TOEICスピーキングテストはスピーキング力を測るテストとして、良いテストだと思いますか?
1) 強くそう思う 2) そう思う 3) あまりそう思わない 4) まったくそう思わない
Q9. TOEICスピーキングテストやTOEICリスニング・リーディングテストなどのテストを受けることにどう思いますか？(複数回答可)
1) 自分の英語力を知るのに役に立つ 2) 英語学習に対するやる気が出る 3) 英語学習に対するやる気がなくなる 4) あまり意味がない 5) やりがいがある

Note. 3) どちらとも言えない in Q4 and Q5 was used only in the 2013 version of the questionnaire. Q6, Q7, and Q9 were used only in 2015 and 2016. Q8 was not included in the 2013 version of the questionnaire.

Appendix C
TOEIC Speaking Test Questionnaire (English Translation)

Q1. How did you find the TOEIC Speaking test as compared to the TOEIC Listening and Reading test (the regular TOEIC)? (Multiple answers allowed)
1) Easy 2) Difficult 3) Undemanding 4) Demanding 5) Fun 6) Boring
Q2. Which do you like better, the TOEIC Speaking test or the TOEIC Listening and Reading test (the regular TOEIC)?
1) TOEIC Speaking test 2) TOEIC Listening and Reading test 3) I like both equally. 4) I don’t like either.
Q3. Were you at a loss what to do during the TOEIC Speaking test due to being unfamiliar with its format?
1) Many times. 2) A few times. 3) Not at all.
Q4. Do you want to take the TOEIC Speaking test again?
1) Very much so. 2) Yes, I do. 3) I’m not sure. 4) Not very much. 5) Not at all.
Q5. Do you think your score will improve if you take the TOEIC Speaking test again?
1) Very much so. 2) Yes, I do. 3) I’m not sure. 4) Not very much. 5) Not at all.
Q6. How well were you able to demonstrate your speaking ability in the test?
1) Perfectly. 2) Quite well. 3) Not very well. 4) Not well at all.
Q7. Compared to speaking directly to a person, do you feel that words do not come as naturally as when speaking to a computer?
1) Very much so. 2) Yes, I do. 3) Not really. 4) No, not at all.
Q8. Do you think the TOEIC Speaking test is good for measuring speaking ability?
1) Very much so. 2) Yes, I do. 3) Not really. 4) No, not at all.
Q9. What do you think of taking tests such as the TOEIC Speaking test and TOEIC Listening and Reading test? (Multiple answers allowed)
1) They are useful for assessing my English level. 2) They motivate me to study English. 3) They make me less motivated. 4) They are meaningless. 5) They are challenging and therefore worth trying.

Note. 3) I’m not sure in Q4 and Q5 was used only in the 2013 version of the questionnaire. Q6, Q7, and Q9 were used only in 2015 and 2016. Q8 was not included in the 2013 version of the questionnaire.