

---

## *Essays*

---

### **Can TOEIC Connect Reception with Production in English?**

---

MICHAEL HERRIMAN

*TOEIC is purported to be a test of communicative English as used in an international context. Its widespread acceptance, and some might claim over-use, particularly in East Asia, leads to the question of what it actually tests. More particularly, we must ask how a test, which is based solely on multiple-choice item responses, can predict the ability to use communicative skills productively in speech and writing. In addressing this question, the issue of what constitutes English for international communication will also be addressed. This will also implicate the issue of what constitutes communicative competence generally. I will look at the connection between receptive and productive skills in language and will also examine the general structure and item types in TOEIC to try to ascertain its appropriateness and utility for the population for whom it is intended.*

#### **Introduction**

TOEIC (the acronym for the Test of English for International Communication) has become a widely used and respected test internationally. As its name suggests, it is intended as a test of communicative English. The fact that it is administered to more than 2 million persons around the world each year indicates its acceptability. The further fact that it was developed and is administered by the Educational Testing Service (ETS, Princeton, New Jersey) gives it a kind of hallmark – one can assume that its items are both reliable and valid, meaning that each item on the test discriminates between good and poor test-takers, and is in itself a valid test of the skill it purports to test. The answer to the question posed by my title might seem obvious; TOEIC is recognized widely as a test of English communication. This answer is not satisfactory however, as TOEIC does not include items involving productive language skills (speaking and writing) and its format does not permit free responses to items on the test since it uses multiple-choice response item types only, requiring no writing.

The term English for international communication is also problematic since it has no immediate denotation. That the use of English is widespread and indeed dominates in international settings is hardly disputable, but it must be asked whether there is any standard against which English can be judged as fulfilling imputed conditions for international communication. The issue is more complicated by facts such as the recognition that English is more widely spoken as a second or other language than it is by first language English speakers, and that varieties of English (known as “Englishes”) have developed locally in many countries such that it may be sensible now only to talk about varieties and deny an international standard to any particular one.

A further general issue is that of what constitutes communicative competence in a language anyway. The concept of communicative competence is a broad one involving not just the comprehension of verbal information, but a range of communicative behaviors that are cultural

and contextual. These will be discussed further below.

The matters just raised seem to widen the gap between TOEIC, with its purely receptive items, and the implicit claim that the test can assess the competence of a person to communicate at an internationally accepted level. Can a connection be made? That is, is there any evidence that would allow the claim that a person with a high score in TOEIC is therefore able to communicate in an international situation? This understanding is in fact a recognition of the motives for developing the test in the first place, as will be seen below.

I will attempt to answer the main question by examining some longitudinal data relating TOEIC scores to scores on other tests of communicative English use. In this study I will use scores from a cohort of more than 200 students taken over a series of tests in one year to see if it is possible to identify more specifically what communicative ability in English is and what TOEIC seems to test. I will first discuss the issue of what might constitute international English communication and then the issue of communicative competence in a language. Then I will discuss the background to the development of TOEIC, its main features and structure, and finally the results of the cohort on 3 TOEIC tests and nearly 20 other tests. It will be seen that there is a surprising degree of consistency in the appearance of a single factor across all tests that might be fairly called communicative ability or competence.

### **International English**

What are the broad parameters of international communication in the English language? Such a question would have been the first consideration of those developing the TOEIC. In test development situations it is common to set up a table of specifications, and for TOEIC this table would have included the set of skills deemed adequate for such communication. Given that TOEIC was developed for use by Japanese speakers to indicate their level of communicative competence in English for business purposes, we can expect that the items of TOEIC have a focus on the context of business. On the surface it appears to focus primarily on such contexts as well as others. The contexts of TOEIC are centered around a corporate culture which includes discussions on the following; meetings and conferences, scheduling, office procedures (ordering supplies, correspondence, making appointments), maintenance of plant and equipment, manufacturing, personnel matters (hiring, firing, promotions), travel and entertainment.<sup>1</sup> These contexts are not specialized ones, however, and there is little of the vocabulary that one would anticipate in the needs of international communication in science, technology, diplomacy, medicine or digital communications where special terminology and higher level registers are involved. It is also fair to say that the registers of TOEIC are those of speech and not writing, and that the examples of written text encountered in the test could not begin to reach that of the business section of national newspapers (in style or depth of analysis, and much less so in length).

The business contexts of TOEIC can be said to be validly international, though that at a low level. The more pertinent question for this paper is that of the extent to which the test can measure communicative competence. This is the issue already referred to as the problem of inferring production from reception. Oral production as opposed to reception involves not just implementing a range of vocal skills appropriate to the target language, but also a capacity to initiate or continue conversations that generally are open-ended. The vocal skills in question include the ability to produce near approximations to the vowel sounds and consonants of the target language as well as the appropriate use of stress, intonation, timing, rhythm (including supra-segmental pausing) and voicing (of consonants). The capacity to carry on a conversation is perhaps even more important to

acceptance into the partnership of a speech situation. Aside from vocabulary and the linguistic skills just discussed, it involves what is usually called communicative competence.

### **Communicative Competence**

The concept of communicative competence has come to play an important part in defining levels of language use in the last three or four decades. It has served to draw attention to the extra-linguistic features of communication particularly, suggesting that a communicatively competent person not only knows the vocabulary and grammar of a language, but also knows how to use it in the variety of situations in which language is called for, and can produce and interpret utterances appropriate to their context of use. The term is more broadly subsumed in the study of pragmatics where a basic assumption is that knowledge of the social aspects of a speech situation probably directs the subsequent linguistic formulation. This means that the speaker needs to understand the social and epistemological setting of a speech situation and know when and how to speak, what not to say, how to continue, change and end topics, understand idioms, agree and disagree (politely or otherwise), and what is taboo, at the least. It also involves knowledge of the extralinguistic and gestural conventions of the speech community in question. Such knowledge is acquired naturally in the speech community of one's first language, but is difficult to acquire in another language. Communicative competence is obviously important in the context of English for international communication, mainly because those who find themselves in international speech situations usually require and expect accurate and unambiguous information exchange. The English language has a high level of redundancy such that meaning or intention can often be recovered despite one or more grammatical or phonological mistakes in utterances. Communicative errors though, are much less comprehensible or forgivable.

### **The Background to TOEIC**

The TOEIC was developed in 1979 in response to the dissatisfaction of industrial bodies in Japan over the existing test of English competency, the Eiken or "Step Test". That test had been supported by the Japanese Ministry of Education (Monbusho), but employers who used it to certify the English language competency of their staff found that even those who had passed Step at the highest level had difficulty in communicating when confronting other English speakers. Representatives of industry in Japan therefore approached ETS (The Educational Testing Service of Princeton, New Jersey) requesting them to develop a test to assess more accurately the level of English competency of staff working in areas where contact with English speakers was required.<sup>2</sup> A steering committee was set up in Japan to handle the test's administration. The then Japanese Ministry of Trade and Industry (MITI) recognized the test, which was then transferred to the auspices of a non-profit organization—the Institute for International Business Communication (IIBC). A high level international advisory board was appointed to oversee the test. In the 25 years since its development TOEIC has become accepted throughout the world. The TOEIC headquarters have subsequently moved to ETS in New Jersey.

From its use in Japan only, the test has spread rapidly overseas. At present, more than 2 million persons take the test each year. About half of these are from Japan, with South Koreans being the next largest group of takers.<sup>3</sup> Recently there has come to be a much wider acceptance of TOEIC in Western Europe and Eastern Europe. There are now more than 50,000 takers in France, where many corporations require employees to certify their English levels via TOEIC. The other

major development has been the penetration of TOEIC into educational institutions, especially in Japan and France. Though intended as a tool for industry's assessments of communicative English, universities and even high schools are now using TOEIC widely to assess students' progress in English and implicitly, as a means of external validation of their programs. The Japanese Ministry of Education has recently set target levels of TOEIC achievement for teachers of English in its schools.

### **The Structure of TOEIC**

TOEIC is a multiple choice item test with two subsections; one designated "listening" and the other "reading". Partial scores are given for each section and these are summed to provide the TOEIC score. The listening section consists of 100 items in 4 different stimulus modes. The first 20 items require takers to look at 20 pictures, one at a time, while listening to 4 statements (designated A, B, C & D) about each one. The statement most closely matching the details depicted is the target. This section can be said to identify the taker's command of canonical statement forms in English. The subjects and objects in the sentence are variously nouns, pronouns and quantified forms. The statements usually contain adverbial phrases of location.

The second set of listening items consists of 30 spoken questions with a choice of 3 spoken answers (A, B & C) for each. Again the test taker has to identify the appropriate answer. The questions generally represent the variety of question forms in English (mainly wh- and verb-auxiliary forms) and rely for their answers on correctly identifying pronominal and verb tense forms used there. The third set of the listening requires takers to listen to a conversation between 2 people and read a question regarding its topic, content (details), or other matters such as inferences that can be made from it. Four answers (A, B, C & D) are given for each question. There are 30 items in this set. The final set of 20 items in the listening section comprises a series of 7-10 short talks. For each talk there are at least 2 questions and 4 possible answers for each, written in the test booklet. The topics vary greatly, but all are communicative inasmuch as the speaker on the tape is heard giving information to an audience of some kind. The content might be a speech, a weather report on the radio, a talk by a guide, a public announcement, a warning, a news report, etc..

The other section of the test, referred to as Reading, has three sets of stimuli. The first is 40 discrete sentence-length cloze exercises, with 4 choices for each. Most of the items target recognition of grammatical conventions or violations, though vocabulary and usage are also prominently tested in this section. The next set is of 20 sentences in each of which are four underlined words or groups of words. One underlined group in each sentence is incorrect and the target is to identify it. The choices are again A, B, C & D. The final set of items comprises texts for 40 questions in reading comprehension. Each text may have 2 to 4 questions related to it. The texts represent normal communicative genres in English and are types not unlike those found in the short talks part of the listening.<sup>4</sup> It might be argued that the second section of TOEIC is mislabeled as reading inasmuch as 60 of the 100 items are discrete grammar or vocabulary tasks, which hardly test knowledge of the conventions of printed texts. The argument that vocabulary recognition and grammar are central requirements of text comprehension, to some extent counters that criticism. In any case most foreign language users of English in communicative contexts might seldom encounter extended prose texts unless they be in technical areas.

### Scoring TOEIC

The test taker is given about 2 hours to answer the 200 questions of TOEIC, the listening taking about 45 minutes of the total and reading allocated 75 minutes exactly. A TOEIC score report contains 2 sub-scores (for listening and reading) and their total which is the TOEIC score. The sub-scores range from 5 to 495 (in 5 point intervals) so the maximum score attainable is 990. The scores are moderated such that it is possible to gain only rough estimates of the taker's raw score (that is, the number of items answered correctly). A TOEIC score is not a categorical score in the way that measuring one's height or weight can be thought to be: it is simply an estimate. The test has a Standard Error of measurement of +/- 30 points, so the score gives an estimate of something between plus and minus 30 points of the stated score. This means that a candidate taking the test successively cannot be said to improve (or have decreased) his or her score unless an improvement (or decrease) of more than 30 points is seen. This is presumably because of the multiple choice format of the test, where nearly all test takers will find they have to guess answers to some items (there is no penalty for incorrect answers). Candidates with low levels of English may find themselves guessing the majority of questions. The algorithm for computing the final score is not made public but it is safe to assume that it compensates for the effects of guessing. It may be assumed also that the items on an official TOEIC are valid and reliable.

### Characteristics of TOEIC

On the face of it, TOEIC can be seen to fairly cover a range of communicative tasks in a wide range of communicative contexts. It exemplifies the language of speech rather than writing and the speech forms represented range from relatively formal (in some of the written texts of part 7) to the idiomatic and colloquial. There is some bias towards North American usage, but not the distinct cultural bias found in TOEFL. It might be said to fairly represent the English mostly used in Japan and perhaps a large part of South and East Asia. The obvious criticism of it is that it is a test of receptivity to the spoken and written language and not of its production. The TOEIC organization has provided estimates to equate TOEIC scores with scores on a number of other tests and to measures of oral competency. It has also produced a table to equate TOEIC scores with commonly recognized levels of English proficiency. The levels are:

905 - 990	International Proficiency
785 - 900	Working Proficiency Plus
605 - 780	Limited Working Proficiency
405 - 600	Elementary Proficiency Plus
255 - 400	Elementary Proficiency
185 - 250	Memorized Proficiency
10 - 180	No Useful Proficiency

It is clear that no attempt is made to equate TOEIC scores with native-speaker proficiency and that only scores above 800~ predict a competence that might be called communicatively adequate. In this paper I will report on some data relevant to this issue.

## This Study

In the present study I have attempted to answer the question posed in the title by comparing students' scores on three IP TOEIC tests<sup>5</sup> with those obtained in sets of other tests also taken twice or three times. My intention was to discover whether there was any relationship between students' performances on TOEIC and on the other tests, and in particular to see if the concept of communicative ability, implied by the title of TOEIC, could be defined more clearly or broken into other components. It is unusual to encounter a population of this size that can provide such a set of data.

## The Subjects

The subjects in the study were 211 Japanese students in the first year of university studies in a faculty of foreign languages. All were intending English language majors. The representation of the sexes was approximately equal<sup>6</sup>. All took the IP TOEIC in the week before their studies began, again at the end of a thirteen week semester and finally at the end of the year<sup>7</sup>. All were studying a common course avowedly in English communication, in this respect differing from many of their age group in other Japanese institutions where studying English might involve studies in literature and linguistics. All were enrolled in 7 courses in English, taken for 27 weeks over the 2 semesters. Each course met once a week for an average of 95 minutes. Over the academic year this meant exposure to about 300 hours of English instruction. Their mean level on entry might be classified as below lower intermediate to elementary, with some having had little or no opportunity to exercise communicative use of the language previously. This is to some extent mirrored in their initial TOEIC score (see Table 1 and Figure 1).

**Table 1: Means etc. for the 3 IP TOEIC tests taken in March, June and December**

### X<sub>1</sub> : TOIEC 1

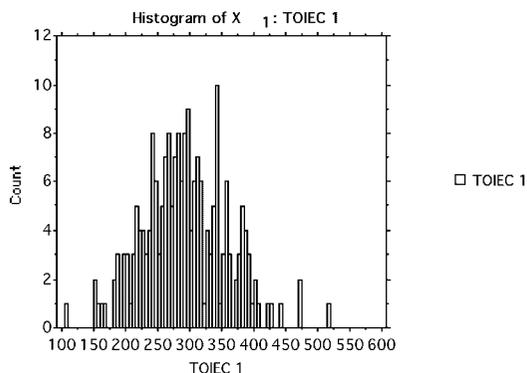
Mean:	Std. Dev.:	Std. Error:	Variance:	Coef. Var.:	Count:
290.121	65.224	4.533	4254.233	22.482	207
Minimum:	Maximum:	Range:	Sum:	Sum of Sqr.:	# Missing:
105	515	410	60055	18299575	2

### X<sub>2</sub> : TOIEC 2

Mean:	Std. Dev.:	Std. Error:	Variance:	Coef. Var.:	Count:
371.011	85.59	6.242	7325.711	23.07	188
Minimum:	Maximum:	Range:	Sum:	Sum of Sqr.:	# Missing:
170	625	455	69750	27247900	21

### X<sub>3</sub> : TOIEC 3

Mean:	Std. Dev.:	Std. Error:	Variance:	Coef. Var.:	Count:
400.604	83.983	6.225	7053.224	20.964	182
Minimum:	Maximum:	Range:	Sum:	Sum of Sqr.:	# Missing:
210	630	420	72910	30484700	27

**Figure 1: Distributional characteristics of the population on the March TOEIC (mean= 290)**

### The Other Tests

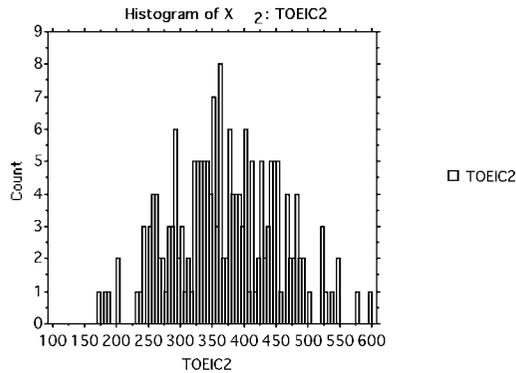
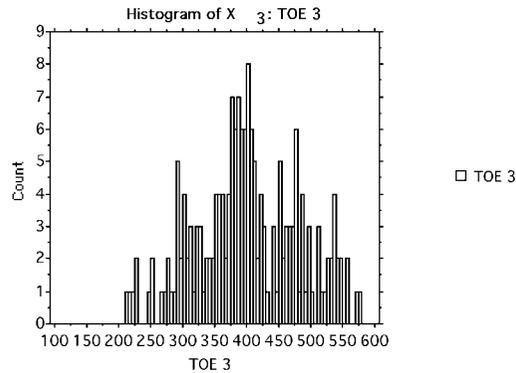
All of the students followed a common course of studies which involved classes designated as follows: English Reading; English Conversation; English Communication Skills (a course in speaking); English Listening; English for TOEIC; English Writing; and Computer Applications. English was used as the medium of instruction in all classes. Exams in the first five named classes constituted the other data used in this study.

The students were taught in classes varying in size from 20 to 38. A common curriculum was followed in every subject using common set texts. No streaming occurred. A common exam was given in 5 of the six classes after 8 weeks of instruction and results, when reported here, are referred to as “mid-semester test”. At the end of the first semester, fifteen weeks after entry to the university, and the end of the second semester, final common exams were administered to calculate the students’ grades for the classes. All exams were based on the material and texts used in classes. The results of these exams, in the case of the first five named subjects, along with the TOEIC scores then constitute the main data of the study. In the exams (mid-semester and end of semester) most questions required simple responses: these ranged from Yes/No, True/False, to single word, to multiple word phrases to sentence length answers.

The data set is then as follows: Scores on 3 IP TOEIC tests (for Listening, Reading and their combined, i.e., TOEIC score) taken in March, June and December; and the mid- semester and final semesters’ test scores on the subjects mentioned above. For most students therefore, there are at least 19 data points.

### Analysis of the Data: Descriptive Statistics

The means and distributional characteristics of the TOEIC tests can be seen in Table 1 and Figures 1, 2 & 3. Means are not given for all other scores for reasons of space, but in most cases the distributions follow a normal pattern with some slight positive skewedness in the first TOEIC distributions. The improvement in TOEIC scores over two semesters (8 months – 300 hours) of study is noticeable. The segment of the population improving is clearly the middle ranking group and those above it. Listening scores show a greater improvement than those designated as reading.

**Figure 2: Distributional characteristics of the population on the June TOEIC (mean= 371)****Figure 3: Distributional characteristics of the population on the December TOEIC (mean= 400)**

### Analysis of the Data: Comparative Statistics

The data were initially compared by correlating the students' scores on all tests. The correlation table (Table 2) shows a pattern of moderate correlations between nearly all tests (the exceptions being those involving the listening component of the first TOEIC test and some of the earlier tests.). The pattern of correlations becomes stronger between the later tests, indicating perhaps the benefits of practice and a coalescence of skills.

**Table 2: Correlation table for the 19 main variables in the study**

	TOEIC1 List	TOEIC1 Read	Total TOEIC1	TOEIC2 List	TOEIC2 Read	Total TOEIC2	Semes 1 TOEIC	Semes 1 Read	Semes 1 List
TOEIC1 List	1								
TOEIC1 Read	.365	1							
Total TOEIC1	.851	.8	1						
TOEIC 2 List	.369	.468	.502	1					
TOEIC 2 Read	.392	.652	.621	.538	1				
Total TOEIC2	.433	.631	.635	.894	.858	1			
Semes1 TOEIC	.338	.529	.516	.531	.628	.657	1		
Semes1 Read	.332	.559	.53	.682	.668	.77	.478	1	
Semes1 List	.348	.557	.538	.521	.614	.643	.463	.705	1
Semes 1 Conv	.479	.524	.604	.57	.622	.678	.501	.677	.661
Semes 1 Comm	.491	.324	.499	.496	.391	.509	.329	.589	.509
Semes 2 Read	.279	.503	.464	.563	.622	.673	.515	.723	.627
Semes 2 List	.434	.637	.639	.56	.632	.677	.499	.661	.813
Semes 2 Conv	.468	.501	.585	.62	.615	.704	.523	.698	.699
Semes 2 Comm	.404	.489	.537	.613	.565	.673	.492	.643	.612
Total TOEIC3	.462	.566	.618	.655	.706	.774	.613	.653	.617
TOEIC3 List	.436	.433	.526	.595	.544	.651	.476	.526	.467
TOEIC3 Read	.373	.558	.556	.552	.693	.704	.597	.618	.614

	Semes 1 Conv	Semes 1 Comm	Semes 2 Read	Semes 2 List	Semes 2 Conv	Semes 2 Comm	Total TOEIC3	TOEIC3 List	TOEIC3 Read
Semes 1 Conv	1								
Semes 1 Comm	.509	1							
Semes 2 Read	.53	.446	1						
Semes 2 List	.787	.541	.694	1					
Semes 2 Conv	.801	.559	.649	.791	1				
Semes 2 Comm	.68	.654	.586	.748	.69	1			
Total TOEIC3	.734	.512	.655	.725	.779	.737	1		
TOEIC3 List	.623	.485	.536	.615	.654	.69	.875	1	
TOEIC3 Read	.662	.411	.612	.654	.71	.6	.876	.533	1

The strong overall correlational pattern suggested the use of factor analytical methods to see if there were some common factors represented in the tests. A principal components analysis was undertaken (omitting the mid-semester tests for reasons of table size)<sup>8</sup> and the main results can be seen in Tables 3 & 4. Four factors fall out (seen in the orthogonal transformation solution – Table 4), but only one is significant (having an Eigen value above 1.0, in this case 11.1 – see Table 3). This factor is represented in all tests which is quite an unexpected result. Its highest loading is on the second and third TOEIC tests, but both the internal tests of listening and conversation load highly on to it, as do all the later tests. The lowest loading (0.57) is seen in the listening component of the first TOEIC test, the component that was least correlated with all other tests<sup>9</sup>. The second factor was represented only in one test (the listening component of the first TOEIC test). The third and fourth factor were not significant. In simple terms, the analysis suggests that underlying all tests is one main component only. The question of what constitutes the factor is fertile ground for speculation. Some speculations are found in the discussion below.

**Table 3: Eigen values for the principal factors identified in the factor analysis**

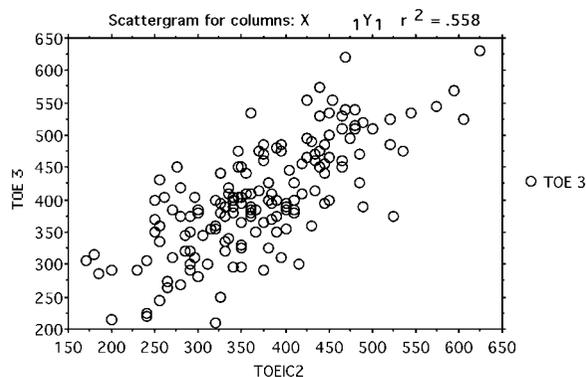
Eigenvalues and proportion of original variance

	Magnitude	Variance Prop.
Value 1	11.111	.617
Value 2	1.199	.067
Value 3	1.031	.057
Value 4	.823	.046
Value 5	.722	.04
Value 6	.49	.027
Value 7	.476	.026
Value 8	.435	.024
Value 9	.39	.022

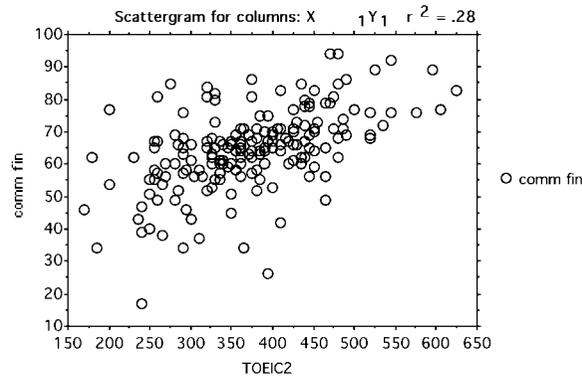
**Table 4: Orthogonal Transformation Solution-Varimax, showing the factor weightings of the 4 principal factors**

	Factor 1	Factor 2	Factor 3	Factor 4
TOEIC1 List	.57	.748	.101	.09
TOEIC1 Read	.715	.193	-.444	-.244
Total TOEIC1	.771	.592	-.185	-.08
TOEIC 2 List	.764	-.149	.027	.279
TOEIC 2 Read	.811	-.071	-.349	.004
Total TOEIC2	.896	-.128	-.169	.172
Semes1 TOEIC	.685	-.052	-.379	.236
Semes1 Read	.821	-.213	.062	-.181
Semes1 List	.786	-.13	.06	-.428
Semes 1 Conv	.831	.013	.145	-.084
Semes 1 Comm	.649	.166	.508	-.055
Semes 2 Read	.762	-.286	-.032	-.144
Semes 2 List	.864	-.037	.113	-.293
Semes 2 Conv	.86	-.068	.184	-.056
Semes 2 Comm	.813	-.061	.302	.052
Total TOEIC3	.896	-.08	.026	.274
TOEIC3 List	.759	-.005	.207	.41
TOEIC3 Read	.81	-.135	-.161	.07

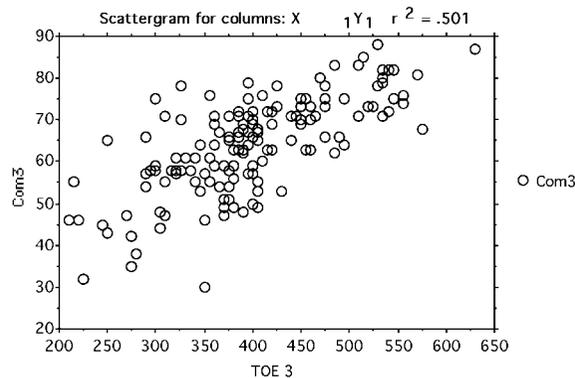
Further comparative relationships pertinent to the main question are seen in scatter-plot for 2 sets of TOEIC scores, and for TOEIC with the communication (speaking) tests (Figs. 4-6 below). The existence of some outliers above and below the central trend is to be expected in a population of this size. A central trend is clear however, showing a strong relationship between the TOEIC 3 total and the score on the last speaking exam (Fig. 6).

**Figure 4: Scatter-plot for scores on TOEIC 2 and TOEIC 3**

**Figure 5: Scatter-plot for TOEIC 2 score and Communication Skills Semester 1 final exam scores**



**Figure 6: Scatter-plot for TOEIC 3 and the Semester 2 Communication Skills final exam scores**



## Discussion

These results are surprising given the varied nature of the courses of study undertaken and the examination formats used. At the observational level, the courses studied by the students and the exams they took apparently focused on the skills stated in the curriculum and only 2 courses used the same text, though for different purposes.<sup>10</sup> What we can say is that the analysis shows the complex nature of both teaching and testing, and by implication the artificiality of breaking the learning or testing process into discrete pieces (labeled the skills of listening, speaking, writing, reading and pragmatics). This is more evident when we consider that such matters as vocabulary knowledge, knowledge of grammar, and the conventions of oral and written discourse probably overlay and intersect all of those skills. To take a skeptical position, it may be claimed that we did not know what we were testing at all. The factor analysis indicates the existence of a singular factor that, since it crosses all areas tested here, might be called a general communicative ability.

But there is no strong warrant for this claim. It might for example, represent vocabulary knowledge; at least we have no obvious grounds for rejecting that or perhaps other possibilities.

Does the result of the analysis advance the initial question of what TOEIC tests? I believe it does, though it is no great surprise to conclude that TOEIC tests the factor that was so clearly represented in all the tests. In the lack of any further information then it may be concluded, albeit tentatively, that TOEIC does test communicative ability.

The fact that the principal factor has a high representation in the test of speaking (the Communication course) suggests that TOEIC may also be claimed to be a reliable measure of productivity as well as receptivity to language. It also suggests that we may not need to devise tests for every so-called sub-skill of language capability, and that a test that we believe to be a valid test of a purported skill may indeed be testing a much more complicated set of skills. This suggests, as already mentioned, that we may be misguided or wasting effort in breaking the task of teaching and learning into discrete parts that we label speaking, listening etc. Even though we may believe that our focus is on such discrete skills, the learner may be getting a different message, absorbing instead a more complex representation of communicative language use that is represented in all language acts of the teacher in the classroom<sup>11</sup>.

### **Limitations**

It must be borne in mind that these results cover a period of study of nine months only. Such a brief period may not be taken as an example of patterns of performance in foreign language acquisition simply because of its brevity. It is well established that foreign language acquisition tends to follow a stage like growth, so it would be wrong to assume that this glimpse of some measures of language performance is generalizable to the larger process.

From the TOEIC point of view, it must be recognized that the students in the study were scoring at the level of elementary proficiency plus in communicative ability, or below it in some cases. Thus it cannot be said that one is dealing with a population with communicative skills across the spectrum of TOEIC scores, a situation which would allow a more robust investigation of TOEIC's predictive capacity in relation to language production.

### **Conclusion**

These reservations notwithstanding, the research answers the main question positively. The strong correlations between the speaking exams (which are tests of speech in genuine communicative situations) and the TOEIC scores, shows that TOEIC can connect reception with production. We may also conclude that there is a common factor represented in all the tests for which there are data here and that it is probably a general communicative ability. It also appears that TOEIC is a test highly suitable for external validation of progress in courses in communicative English learning, especially in an intensive situation such as that faced by the students in this study. The further question for this research would be that of whether TOEIC scores would correlate with or predict the other area of productive skills, namely, writing. The main question could be further advanced in looking at a population representing a greater range of TOEIC scores. The issue of what constitutes an International or World English deserves much greater attention also.

## Acknowledgements

I would like to thank Bruce Monk and Gerrold Meister for their perceptive reading and comment on the draft of this paper. Any subsequent errors or omissions are my responsibility.

## Endnotes

1. Knowledge of actual items and contexts can only be inferred since the tests (i.e., the multiple forms) themselves are confidential. However, the TOEIC organization publishes examples of items and these concur highly with the many books of practice items published independently. The TOEIC administrators also employ item writers outside of the company, who have contributed to the public knowledge of the content and contexts of a TOEIC test.
2. TOEIC Steering Committee (1991) *Test of English for International Communication: History and Status*
3. In Japan, about half of the test takers take the Institutional Program (IP), a version of the TOEIC test developed especially for use by institutions and administered by and within them. These tests are scored by ETS however and the items on the test remain secure. Scores on the IP are not officially recognized as TOEIC scores, but many industries and educational institutions accept them as valid indicators of a TOEIC score.
4. The texts typically include letters, memos, labels, advertisements, charts, tables, diary entries, schedules, notices, instructions for using products, newspaper articles, etc..
5. The IP TOEIC is the aforementioned form of the test administered in and by an institution, but provided and scored by ETS. It may be regarded as a parallel form of the official TOEIC.
6. In this population the female-to-male ratio was 6:4.
7. The 3 tests were given at the end of March, mid-June and early-December.
8. The result when the mid-semester tests were included, increased the principal factor weighting and Eigen value.
9. In this case it must be remembered that the first TOEIC was taken in their orientation program, before the students had commenced their studies.
10. The Listening and the Conversation courses both used *New Interchange 1*.
11. What I have in mind here is that it is possible to identify at least 3 distinct contexts of language use in the classroom: these might be labeled the interpersonal, the transactional and the focus (in this last case, directed to the English language as an object of study). If the teacher uses English in all contexts, the communicative use of English is, ipso facto, put before the student as a part of the experience of learning. What is significant here is that in the case of most of the students in the study, any previous acquaintance with English was with the language presented in the focus mode only, that is, as an object of study. The Japanese language was normally used in the first two contexts. Consequently most students will have experienced little by way of English in real communicative usage in high school.