

Voices in the Field: An Interview with Kazuhiko Saito

by Mark Chapman

Kazuhiko Saito is the General Manager of the Public Relations Department at the Institute for International Business Communication (IIBC). IIBC's core activity is the administration of the TOEIC®. IIBC is a non-profit organization which has been operating under the auspices of the Japanese Ministry of Economy, Trade and Industry since 1986. This interview was conducted at the IIBC office in Tokyo on May 13th 2004.

Could you briefly outline the history of the TOEIC?

The late Yasuo Kitaoka, a former vice chairman of IIBC, recognized the strong need for a test that measured English-language skills and in 1978 visited Educational Testing Service (ETS). Mr. Kitaoka was looking for a test that would apply to everyday situations involving international communication, and measure not only the communicative English ability of Japanese businessmen, but also their communication skills in English as a universal language. After a series of negotiations, ETS began developing the TOEIC test, and in order to promote the TOEIC a steering committee chaired by Mr. Yaeji Watanabe was established in 1979 within World Economic Information Services (WEIS), a non-profit organization under the auspices of the Japanese Ministry of Trade and Industry.

The first open TOEIC test was administered in Japan in December 1979 and was taken by 2,773 people and the average score was 578. Since then there has been a dramatic growth in the number of people taking the test. Some 3800 corporations across virtually all business fields have recognized the usefulness of the test.

How is the TOEIC administered in Japan?

The TOEIC test is currently administered through the Secure Program (SP) eight times a year. The SP administration uses newly created test forms that are administered on the same date at the same time at all test centers. All answer sheets are sent to a Tokyo computer center where the answers are digitalized and sent to ETS in Princeton for analysis of test questions and scoring. Score data is returned to Tokyo and examinees typically receive their scores one month after administration.

Can you mention something about the IP-TOEIC?

The Institutional Program was started in 1981 and is offered to business firms, universities, language schools and other approved institutions. These tests can be arranged for any date convenient to the institution as long as notice is given ten days in advance of the test date. The same high degree of test security associated with the SP administration is not provided. However, forms are used every six months, and measures are taken to ensure the same examinee does not take the same form twice, even within a six month period.

How many people take the TOEIC every year and what percentage of the total does Japan account for?

In 2003 there were approximately 3,400,000 examinees globally. Japan accounted for 1,423,000 candidates or 41.8% of the total. The number of people taking TOEIC has doubled worldwide since 1999 and in Japan alone more than 11 million people have taken the test since its introduction in 1979.

What changes have you noticed in the profiles of the test takers over the years?

In terms of English-language needs, when overseas business centered on imports and exports, only the personnel at trading companies and in the international business divisions of some manufacturers needed to use English for their everyday work. With time, however, English-language skills were required of managers, technical personnel, and other staff that were dispatched overseas. Major electric appliance manufacturers, as well as some automobile manufacturers and other companies, began to establish training programs to teach these employees English.

As of 2003, the average IP TOEIC score in Japan was 460 and the average SP TOEIC score was 570. These figures are relatively unchanged from 1998. In 2002 an analysis of examinees by gender (in the IP program only) showed that in companies, men made up almost 80% of the total. In schools, the picture is quite different, with female examinees representing around 60% of the total. The majority of examinees in the SP program are in their twenties, whereas for the IP program there are approximately equal numbers of examinees in their twenties and thirties.

What do you think are the reasons for the commercial success of the TOEIC?

The two main reasons are the nature of the test itself and structural changes in the Japanese business situation since 1979. The most important feature of the TOEIC is that the examinee's proficiency is expressed as a numerical score between 10 and 990, rather than simply a pass or fail grading. The scale is constant. If there is no movement in the proficiency of the examinee, his or her score will remain essentially unchanged. This consistency allows examinees to obtain a fairly accurate measurement of their current proficiency in English, and a target score to work towards. Because there is no failing grade, examinees don't lose motivation so easily and if a test has several different grades, companies may need to prepare several versions of the test. This problem is avoided with TOEIC. In addition, TOEIC aims to be a fair evaluation of English communication skills so questions that cannot be answered without a certain cultural background, or expressions unique to the U.S. or UK, are avoided to the greatest extent possible. This means that the examinees' communication skills can be evaluated regardless of nationality or ethnicity.

The success of the TOEIC may also be linked to the globalization of the Japanese business world and the internationalization of Japan. As the Japanese economy grew, Japanese corporations made inroads overseas and there was an inevitable need to dispatch more employees abroad. Companies were faced with an immediate challenge of fostering personnel who could work in overseas environments. In the late 1980's many companies began initiatives to improve the overall English proficiency of their employees through a "bottom-up" process. It was no longer sufficient to train certain employees to work as specialized international personnel. Companies began providing language training not only for new recruits, but also for all of their other employees. Intensive English lessons or English-training camps in which employees were immersed in English became a common feature of the late 1980's and early 1990's.

The 1990's saw the collapse of the Japanese economic bubble and a shift in policy for corporate English-training programs. Companies began to question how their programs were administered, in light of the lack of forthcoming results. Consequently, the quality and effectiveness of the training programs came under greater scrutiny and companies were faced with the need to cut costs. In the mid-1990's, most of the electrical manufacturers in Japan who had previously conducted their own testing programs decided to switch to the TOEIC test to evaluate employees' English-language skills. New recruits and college students were also impacted by the burst of the economic bubble. It forced them to pay greater attention to their careers, which resulted in an increase in the number of TOEIC test-takers among university students. Companies that took TOEIC scores into consideration when hiring increased from 44% in 1989 to 59% in 1998. TOEIC is now offered in one form or another at over 200 two- and four-year universities. In 1999 the Ministry of Education approved the TOEIC test as a way to obtain university credits. By 2003, 236 universities in Japan were using TOEIC as a criteria for university entrance and 266 universities were using the test for accreditation. Hence, the structural changes in the Japanese economy have helped boost the popularity of TOEIC. This has been followed by rising demand for the test among college students. These two factors, combined with the nature of the test are responsible for the commercial success of TOEIC.

Can you explain the statistical procedures behind the “equating” process?

The process by which scaled scores are made equal, sometimes referred to as the equating process, makes adjustments to raw scores from test forms of different difficulty levels. In this way, the same level of performance in Listening Comprehension (or Reading Comprehension) is reported as the same scale score regardless of the specific TOEIC test form used. This process assures that each version of the test is as equal in difficulty level as possible to any other version. Otherwise, a person's score on any particular test form would pertain only to that test form and would not be a valid indicator of their English ability as compared to other people who took other forms of the test.

Equating determines the number of items a candidate needs to answer correctly on any particular test form in order to acquire a particular scale score on that test form. For example, a candidate taking one version of the test may need to answer 88 items correctly to receive a scale score of 450, while a candidate taking another version of the test may need to answer only 85 items correctly to receive the same scale score. This indicates that the first candidate's test form was easier than the second candidate's test form.

Each TOEIC test form is equated back to two older TOEIC test forms by incorporating a block of items from each old form in the new test form. The equators for each TOEIC test form are chosen by test developers based upon item reliability (r-biserials and delta values) and upon the test specifications. A series of computations are used to equate the test forms. The equating computations are applied to a hypothetical sample, known as the “equalized group.” The equalized group contains two Japanese subgroups from the first secure administration of new test forms, those candidates who are affiliated with a company (Affiliates) and those who are not affiliated with a company (Non-Affiliates). The relative number of Affiliates and Non-Affiliates changes with each administration, hence, the groups must be equalized in order to properly equate. TOEIC test scores are reported on a scale which was instituted on the first TOEIC test administration.

Which linguistic skills do you consider TOEIC to be a valid measure of?

TOEIC is designed to provide information on an individual's overall communication ability, including speaking and writing skills. Although only listening and reading skills are directly measured, solid statistical evidence suggests there is a significant correlation between passive skills and active skills, particularly between listening and speaking. A series of validation studies were performed to determine what extent TOEIC test scores could predict how well the examinee could speak English. From the results of these validity studies, we can predict fairly well how a person would be able to perform in English, and what range of competence that person has been able to attain in each of the language skills.

Would you recommend that Japanese corporations use only TOEIC in making linguistic decisions about employees? Is it necessary for corporations to employ more direct tests of speaking and writing (in addition to TOEIC) to gain a really accurate assessment of individual linguistic proficiency?

We do not recommend corporations use TOEIC in isolation. However, this decision depends on the situation each company finds itself in. Corporations are already very familiar with the TOEIC test and the information it provides, so if they feel they need additional information about speaking or writing ability, they can adopt separate tests of speaking and writing. We do offer the Language Proficiency Interview (LPI) for examinees who have scored 730 or higher on the TOEIC SP test. The LPI is a 20-25 minute conversational, face-to-face interview conducted by a certified interviewer.

We believe it would be good to have speaking and writing tests. However, in many companies listening and reading are the most important skills for employees. The increasing importance of e-mail and the internet means that reading is an essential skill for employees of Japanese corporations. Listening comprehension is also a requirement given the number of international telephone calls and teleconferences. Having said that, if we could find an easy and cheap way of measuring writing and speaking, we would like to do so. Given the very large numbers of people taking TOEIC it is difficult to test speaking and writing.

Would you describe the TOEIC as a reliable progress and/or diagnostic test?

The TOEIC is recommended for use as a progress test, but we cannot say it is a diagnostic test. Many companies utilize TOEIC as a way of measuring employees' proficiency gains after English training courses. TOEIC scores can also be used to determine the effectiveness of English-language training programs.

In 2005, TOEFL will be re-launched with a speaking element, along with the Test of Written English. Are there any plans to change the TOEIC by introducing a compulsory writing or speaking section?

At the present time there are no plans to introduce a compulsory writing or speaking section. We are considering and studying speaking and writing tests, but we have no official plans to introduce a new test.