The 150-Year History of English Language Assessment at Schools in Japan

(Running head: 150-YEAR HISTORY OF SCHOOL-BASED EFL ASSESSMENT IN JAPAN)

by

Miyuki Sasaki
Faculty of Foreign Studies,
Nagoya Gakuin University

Abstract

In the present study I describe the 150-year history of school-based English language assessment in Japan. This history is divided into four major periods according to the purposes of English language education set by the government in the different periods: (1) 1860 to 1945, when English was first introduced and taught in schools mainly for elite classes; (2) 1945 to 1960's, when English became part of the compulsory education for the first time; (3) 1970's to 1980's when English began to be regarded as the most effective means to communicate with foreign people in the rapidly shrinking world; and (4) 1990 to the present, when several innovative policies have been introduced to the classroom measurement systems. I describe how assessment practices for English education at public schools in each of these periods were and have been affected by various factors, including political, economic, and demographic changes in society, as well as academic paradigm shifts in the fields of education and
applied linguistics. On the basis of the historical analysis of such changes, I also propose possible future studies that may provide useful suggestions for successful execution of innovative policies for language testing. (192 words)
On January 21st, 2006, a total of 492,596 students (about 40% of the high school graduates of that year) took the English test given by the National Center for University Entrance Examination in Japan. The test was necessary for university admission for the following academic year (starting in April, 2006). This year’s English test was memorable in that the Center added a 50-point listening comprehension section to the original 200-point pencil-and-paper section for the first time in its 26-year history. Up to this time, the examinees’ listening ability was measured only indirectly through written items of accent-location and pronunciation. Because it is such a high-stake test taken by so many examinees, the Center has been extremely careful in deciding to include the listening section in spite of the long and well-supported public criticism that their English test was invalid in terms of measuring examinees' oral skills (e.g., Negishi, 1997; Saito, 2002). The introduction of the listening section had been publicly announced for four years (since July, 2002), and when the test was actually given, an individual disposable IC player with an earphone set was distributed to every examinee for the purpose of providing them with the most equally fair test-taking conditions.¹

Regarding entrance examinations, this was only one of the many innovational measures that the Japanese government has recently taken. Other steps include repeated public appeals that universities use not only a single test score, but also other multiple measures from various perspectives (e.g., reports from high schools, essay examinations, interviews), which should be given on multiple occasions to ease the excessive competition (the Central Council for Education, 1997, 1999). For English tests, for example, the Central Council for Education even advised the Ministry of Education that English tests should be removed from the entrance Examinations for
some universities that have long required a certain level of English proficiency of the applicants, while suggesting the possibility of using their high school English grades or their scores from other commercial English tests instead (the Central Council for Education, 1997). How has the Japanese government come to arrive at the present attitude toward school-based English assessment? What factors have influenced their policies? These questions have motivated the present study. In order to properly understand the current situation, and to explore more "democratic assessment" (Shohamy, 2001, p. 373) for the future of English language assessment, it is useful to examine the historical context.

In the present study, I describe the 150-year history of school-based English language assessment in Japan. I used the term "assessment" instead of "testing" here, following Lynch's (2001, p. 358) definition of assessment as "the systematic gathering of information for the purposes of making decisions or judgments about individuals" (p. 358), which includes not only conventional tests but also other types of assessment such as school grades and principal recommendations. Because the scope permitted by the journal-paper length is quite limited, I mainly focus on the changes in English tests used for school admission (i.e., entrance Examination), because in Japan people have treated them most seriously among all English tests since English started to be taught 150 years ago (Imura, 2003; Kawazumi, 1999; Watanabe, 1995). Other types of language assessment such as English language tests given outside schools and assessment in other languages are beyond the scope of the present study. In terms of method, I basically followed Isaac and Michael's (1981) framework for historical research: The purpose of the present study was thus to "reconstruct the past objectively and accurately, often in relation to the tenability of an hypothesis" (p. 42). As working hypotheses, I employed the following two:
(1) The history of English language assessment at schools in Japan can be categorized according to the intended goals and the degree of popularization of school-based English education in Japan.

(2) English language assessment at schools in Japan has been influenced in different ways by socio-political, economic, academic, and historical/cultural factors in different periods.

When collecting data, I constantly checked the validity of the data by asking the two questions recommended by Isaac and Michael (p. 45) for conducting a rigorous historical study: the 'external criticism which asks, "Is the document or relic authentic?"' and the 'internal criticism which asks, "if authentic, are the data accurate and relevant?"' (see also Matsuda, 2005). However, readers are still warned that the data for the present study consisted of what Connors (1992, p. 15) called the "three elements: the historian's perceptions of the present, her assemblage of claims based on study of materials from the past, and an ongoing internal dialogue about cultural preconceptions and prejudices and the historian's own." That is, my writing in this paper was influenced by my having been an insider in the system as a student (1963 to 1987) and as a teacher (1991 to the present).

Japanese people began learning English in 1808 when the British His Majesty's Ship Phaeton appeared in the Gulf of Nagasaki. Although Japan was closed to foreign countries except for the Netherlands and China, this ship forced its way into the Gulf by first pretending that it was Dutch, and later threatening to set fire on the city of Nagasaki (Imura, 2003). Being forced by such a foreign power, the Tokugawa Shogunate government ordered their Dutch TSUUJIs, Japanese officials who used Dutch for diplomatic and trading exchanges between Japan and the Netherlands, to learn English to prepare for further threats to similar sorts. However, formal English language teaching/learning at schools did not begin until 1860 when English joined the
previously taught Dutch at BANSHO-SHIRABESHO, the first public foreign language school, which had been founded by the Tokugawa Shogunate three years earlier (Kawazumi, 1978). The present study describes the history of English language assessment for about 150 years since then. I divided the 150 years into four periods according to the intended goals and the degree of popularization of school-based English education in Japan of the different periods.

PERIOD 1 (1860 to 1945): ENGLISH EDUCATION FOR THE ELITE

From 1860, when the BANSHO-SHIRABESHO started to teach English for the first time in Japan, till 1945 when Japan lost the second World War against the Allied Powers, English was taught mainly to children from the middle to upper classes. BANSHO-SHIRABESHO was only for a limited number of high-class SAMURAI warriors of the Tokugawa Shogunate. It started as a school for the TOKUGAWA Shogunate, but it continued to serve as the major public foreign language school even after 1867, when the new Era of Meiji started as the result of the last Shogun Yoshinobu's returning the reins of government to Emperor Mutsuhito. Because the Meiji government strongly hoped to modernize the country by absorbing new knowledge introduced by European countries and the United States, it often sent students to, and hired teachers from these countries (Kawazumi, 1978). Consequently, the French and German languages were also employed as a means of such modernization in the beginning, but the Meiji Government soon decided to make the language of instruction at the national university (originally the BANSHO-SHIRABESHO mentioned above) English only, adopting the policy of "(Eigo-Honni-Sei) English-as-the-standard." Thus, for the first 20 years of the Meiji Period, English was not the target of learning, but the means of acquiring other new skills and knowledge. After the first 20 years, however, a sufficient number of the textbooks
used at the university level were translated into Japanese, and English quickly became a foreign language only to be taught as a subject at school.

In 1872, the Meiji government promulgated the first educational law, and planned to build 8 universities, 256 middle schools, and 53760 elementary schools in the future. Although this plan did not become realized very soon (e.g., there was only one university for the following 20 years), in 1886, the government enacted another law requiring that every Japanese child receive at least four years of compulsory education at elementary school. The duration of compulsory education was extended to six years in 1907, and elementary school attendance reached 98% in 1911. However, English was not a required subject for most Japanese elementary school pupils although it was taught in some (about 8.5% of all elementary schools in 1940, for example) private and higher-elementary schools (attended after the regular 4-year or six-year long elementary schools, see Imura, 2003). In most cases, students started to study English after they entered middle schools (e.g., junior-high schools, junior-high girls school, vocational schools) around the age of 10 to 12 years old. Going to those five-year-long middle schools, which were supposed to prepare the students for still higher education, was not compulsory, and only about 20% of elementary school graduates could go on to receive such or higher education until after 1945.

As I mentioned above, Japan spent only twenty years or so to change the status of English from a second language to be used as a means of communication at school, to a foreign language to be studied as a school subject. This was because many people thought that education through a second language was like a humiliating "colonized state (translation by M. Sasaki)" (Imura, 2003, p. 85). This attitude went as far as banning instruction in any foreign language at the (only national) university in 1882 (Kawazumi, 1999). Such a change of status for English consequently determined the
nature of its instruction at school during this period and since then. Because it was
difficult to travel abroad during this period, very few English teachers had the chance to
practice their English for communicative purposes (Suzuki & Wakabayashi, 1999).
They consequently taught English as they had been taught, mainly explaining its
grammatical structures in Japanese, and translating the given English texts into Japanese.
This "grammar-translation" method was useful when the teachers were "not competent
enough to give their students the oral skills" (Tajima, 1978, p. 225). Furthermore, this
method had successfully been used for teaching Chinese as a written form ever since the
language was introduced around the fourth century (e.g., 61 characters carved on the
sword at Isonokami Temple, Nara in 369 A.D.). Watanabe (1995) maintained that
such a method was useful for "cultivating" the learners' mind by forcing them to analyze
the differences between Chinese and their mother tongue, Japanese. Cultivating
students' minds subsequently became one of the major purposes of teaching EFL in
Japan, although the Ministry of Education set the purpose of foreign language
instruction to "understand and use" the language as early as in 1901 through the
Enforcement Regulations for Middle Schools.

The fact that English was taught to cultivate students' minds rather than as a means
of communication also determined the content and form of English tests used in Period 1.
Such an attitude was especially reflected in the entrance examinations to upper-
secondary (or pre-university) schools. Those examinations usually consisted of items
requiring translation of syntactically/semantically complicated sentences from English
to Japanese, and vice versa (Nagahara, 1936 cited in Imura, 2003). The texts used for
such items were often taken from British and/or American literary works with moral
lessons, which were thought to be appropriate to "cultivate" students' mind (e.g., The
Use of Life by Avebury, 1894, Ethics for Young People by Everett, 1893, Self-Help by
Smiles, 1859; see Ikeda, 1967; Imura, 2003 for more details). English studied for this particular purpose was (and still is) called "JUKEN EIGO (entrance examination English)." Because the competition to be admitted to upper-secondary schools was quite severe (e.g., see Kume, 1918), studying JUKEN-EIGO effectively became the major purpose of middle school English classes (Nagahara, 1936 cited in Imura, 2003). Reference books for student preparation for such entrance examinations appeared as early as in 1905 (Professor Nannichi’s How to translated English sentences published by Yuhoodo, Tokyo), and sold very well.

The English education of Period I, which was not very practical because it was taught mainly for preparing students for the entrance examinations, was sometimes severely criticized (e.g., Fujimura, 1927), but continued to survive even during the second world war conducted against the two English-speaking countries of the United Kingdom and the United States (Kawazumi, 1999).

PERIOD 2 (1945 to 1970): ENGLISH EDUCATION FOR EVERYONE

On August 15th, 1945, Japan surrendered, ending the second world war, and was occupied by the General Head Quarters (GHQ) of the Allied Powers until 1952, one year after Japan concluded a peace treaty with the United States and allied nations. Under GHQ's guidance, Japan democratized a number of social systems, including elections, land ownership, and education. Following the American model, the whole school system was changed into a single-track system of six years at elementary school, three years at junior high school, three years at senior high schools, and two to four years at colleges or universities. The new school system took effect in April, 1947, and the first nine years of education became compulsory. English was an elective subject to be taught at junior high schools, but it virtually became a required subject in almost
all junior high schools after it was included as a subject used for the entrance examinations for proceeding into senior high schools, at the end of 1950's.

The most notable difference between this period and the previous one was that now everyone had a chance to learn English as a foreign language (EFL). That is, this period represents the beginning of mass popularization of EFL education in Japan. Ironically, however, this very effort produced many English teachers with poor qualification. In order to extend the length of compulsory education from six to nine years, postwar Japan suffered from a serious shortage of both school buildings and teachers for an extended period. For example, when this new education system started in 1947, it was assessed that the system was short of 83000 junior high school teachers (Yamazumi, 1987), and the subject of English was no exception. To supplement this shortage, teachers of other subjects such as physical education or music, were forced to teach English (or vice versa) (Fujii, 1971).

Various measures were taken to try to solve the problems faced by such new high school English teachers. First, in 1947 and 1951, the *Suggested Course of Study in English for Lower and Upper Secondary Schools* were issued by the Ministry of Education, Science, Sports and Culture (henceforth, the Ministry of Education) with the intention of making them a kind of reference book for those new teachers. To fulfill this purpose, the 1951 version was made much longer (759 pages with 3 volumes) than the subsequent versions, containing not only the objectives of teaching English (they emphasized both practical purposes and cultivation of mind), but also a list of teaching materials, reference books, and even an explanation of English pronunciation and intonation (Imura, 2003). Another measure taken by the Ministry of Education to improve the quality of high school English teachers was their support for the foundation of English Language Education Council (ELEC) in 1956. ELEC propagated Fries'
(1945) Oral Approach based on the ideas from structuralism in linguistics and behaviorism in psychology, which were then prevalent in the United States. Another branch of the audio-lingual method, Palmer's (1921) Oral Method, had already been introduced into Japan as early as in 1922, both by Palmer himself and by the Institute for Research in English Teaching founded by him in 1923. Both Palmer's Oral Method and Fries' Oral Approach can be characterized by their sole emphasis on repeated spoken drills (e.g., pattern practice) especially in the beginning part of the learning process. Believing in their effectiveness, the Ministry strongly advocated these two audio-lingual methods. However, these teaching methods did (or could) not become very popular because "(they) cannot be successfully carried out by the teachers unless the teachers themselves are proficient enough in (speaking) the L2 (but most of them were not)" (Suzuki & Wakabayashi, 1999, p. 10, translated by M. Sasaki), and use of these methods gradually declined by the 1970's.

When we look at the EFL tests made in this period, we can see similarity with those tests made in Period 1 mainly because the test writers, i.e., the English teachers, carried over their prewar teaching beliefs and methods into their postwar classrooms (Fukui, 1975). This is especially true with senior high school teachers and university/college professors. Hoshiyama (1978, p. 10) described the teaching methods of senior high school teachers after World War II as follows:

As most of the teachers have majored in either English literature or American literature in colleges or universities, so they are regarded as qualified, but it is natural for them to teach English in the way they were taught in their school days, namely, the old, grammar-translation method."

The old grammar-translation method was also inherited by university professors, and English entrance Examination that they made mainly focused on reading comprehension
of texts from the classics, such as Bertrand Russell, or literary work such as George Orwell (Imura, 2002). In fact, many popular preparatory reference books for university entrance examinations survived after 1945, and some are still on sale even now (e.g., Sensaku Hara's *A Minute Study of Standard Passages from Eminent Authors*, first published in 1933, has sold over 9.5 million copies, and its fifth edition is currently still on sale). The only differences between the two periods were that postwar examinations used much longer (300 to 500 words) texts, and that the postwar examinations contained not only the "translate from English to Japanese or vice versa" items used in the prewar periods, but also additional reading-comprehension and/or grammatical items, as well as items asking about accent location/pronunciation of written words (Hayama, 1963; Imura, 2003; Koike, 1971; Nakayama, 1950).

In contrast, English entrance examinations for senior high schools became more multiple-skill oriented after English became a virtually compulsory subject at junior high schools around the end of the 1950's (Koike, 1971). The examinations typically had so-called SOGO-MONDAI (comprehensive tests) containing various types of items such as grammatical and comprehension items for a 150 to 300 word text, accent-location/pronunciation of written word items, and writing items. Because all applicants took one common examination if they applied for public senior high school in the same prefecture, the test writers were under the guidance of their prefectural government, and they tended to be careful to faithfully follow the curriculum content suggested by the Courses of Study at the given time, which emphasized practical use of all four skills of listening, speaking, reading and writing. Although the speaking skill was (and has not been) tested in these examinations, some prefectures started to test listening ability in the 1960's.
PERIOD 3 (1970 to 1989): ENGLISH EDUCATION FOR PRACTICAL PURPOSES IN THE RAPID GLOBALIZATION

Period 3 marks the beginning of a qualitative change in English education in Japan. In Periods 1 and 2, English tended to be regarded as a unilateral means of importing foreign culture and knowledge whereas from Periods 3 on, English came to be regarded as a means of communication with other people in the world. This way of looking at the English language was influenced by Japan's economic growth as well as the internationalization of English itself, both of which have influenced the EFL teaching/assessment practices in Japan after Period 3.

First, Japan's economic growth brought about three social changes that had significant impacts on EFL classrooms in Japan: (1) rapid increases of those who proceeded to non-compulsory higher education, (2) an increase of those who traveled/lived abroad, and (3) an increase of English-speaking teachers. After Japan's complete defeat of World War II, its economy was revived in part with the help of the Korean War procurements between 1950 and 1953. From the time Japan's economy continued to grow rapidly, and in 1968 it achieved the second largest Gross National Product among capitalistic nations. As the result of this nationwide economic growth, the mean family income grew rapidly, and more and more parents started to send their children to post-compulsory level schools. In 1948, the percentage of junior high school graduates who entered the senior high schools was about 40%, but only 26 years later in 1974, it exceeded over 90%. Similarly, in 1975, the number of the students who were enrolled in either four-year universities or two-year colleges exceeded two millions, about one third of the college-age population.

Such a popularization of higher education had two serious effects on classrooms in Japan. First, it resulted in many students who could not digest what they studied at
high schools. The Ministry of Education tried to solve this problem by simplifying the content of the junior/senior high school syllabi, and in the Course of Study promulgated in 1977, the number of hours of English classes given at public junior high schools decreased on average from four to three (the total annual classroom hours were also reduced). Another serious problem caused by the rapid popularization of higher education was the severe entrance examination competition. Although the government tried to respond to the need for higher education by building public high schools and by giving permission for the establishment of both public and private universities (e.g., between 1950 and 1975, "the number of four year universities jumped from 201 to 426," Matsuyama, 1978, p. 35), getting into the non-compulsory senior high schools and universities/colleges of one's choice remained highly competitive in Period 3. In order to lower the risk of one-time-only tests for getting into senior-high schools, many prefectures in Period 3 started to recommend that public senior high schools give multiple opportunities to take tests to their applicants, and to consider the three-year junior high school grades and conduct of the students in the selection processes (Imahashi, et al., 1990). Supporting this extreme competition was the Japanese belief that one needs to be admitted to a high-ranking senior high school to pass the entrance examination to get into a high-ranking university, which was supposed to guarantee one's subsequent high-ranking position in the society.

The competition was especially so severe for the university entrance examinations that every year some applicants even committed "suicide mainly because of the pressure and the anxiety of the examinations" (Matsuyama, 1978, pp. 35 - 36). Such applicant anxiety was partly caused by the fact that for selective purposes entrance examinations of many prestigious universities asked questions that were too difficult for many applicants. This had already been a problem toward the end of Period 2. For
example, criticizing the university entrance Examinations given in 1960, Imura (1997, p. 41) pointed out that "in order to pass first-rate universities' entrance examinations in English, applicants needed to know 2000 extra English words that were not included among the 5100 words that had been determined to be learned (by the Ministry of Education) during the six years of high school English education (translated by M. Sasaki)." This problem appears to have continued at least into the 1970's because Matsuyama (1978, p. 41) criticized that some words, such as "admonition" and "incinerator," that appeared in the texts of some prestigious universities' entrance examinations in 1977 were examples that were "beyond the high school students' reach to memorize."

The Ministry of Education's efforts to deal with such problems with the university entrance examinations were met by the government's policy of trying to centrally control Japan's education in Period 3. Motivating this policy were the circumstances that by then Japan had been incorporated into the western anticommunist camp in the Cold War structure between the East and the West after the Korean War (1950 to 1953) (Yamazumi, 1987). In order to carry out this policy, the Japanese government exerted considerable effort to regain its central role in education, which had once weakened after World War II. For example, in 1958, the Ministry of Education decided to add law-binding force to their Course of Study. The ones issued in 1947 and 1951 were only "suggested," but after the 1958 version took effect, all teachers have had to obey the goals set by the Course of Study, and teach the required materials.

In 1970, the Japanese government thus tried to both control and solve the problem of "too-difficult" university entrance examination by proposing the development of a common examination shared by all public universities. The government had actually attempted to give common entrance examinations to university applicants twice in 1948
and in 1963 during Period 2, but these tests were not successful because "almost all the universities intentionally neglected or openly opposed them" (Matsuyama, 1978, p. 42) feeling that their autonomy would be threatened. In Period 3, however, the government's third attempt succeeded partly because the entrance examination problem became so serious that it became necessary to publicly intervene in some way, and partly because this time the Ministry of Education wisely let the university professors themselves take the initiative in developing the common test. Subsequently in 1971, the Association of National Universities agreed to develop this "Common Test." The next year, the National Center for University Entrance Examination (conventionally called DNC) was established to take the role of developing the test, and after several trials with preliminary versions, the first Common Test was given in 1979 to 320,000 applicants of all national and municipal universities.

These applicants had to take all examinations in the five subjects of Japanese, social studies, mathematics, science, and a foreign language (English, German, or French). In order to solve the problem of "too difficult" items, the Common Test was made to measure the core knowledge that high schools students were supposed to have acquired by the time of their senior high school graduation. All items were multiple-choice and the results were processed by computer. The application procedure was somewhat complicated in that after the DNC published the correct answers for all the items, the applicants self-scored their answers, and applied to the universities that were likely to admit them. After they applied to the individual universities, they again had to take the entrance examinations of those universities, but this time in fewer subjects. The answers for these second examinations were often not in multiple-choice form, but in descriptive forms. In contrast to these public university applicants, private university applicants who occupied more than 75% of all university applicants, did not take this
The new Common Test was supposed to serve for multiple purposes such as alleviating the severe university entrance examination competition, providing a standardized measure to assess the university applicants' ability, and normalizing the senior high school English classes. The Test did achieve these goals to some extent, but it also brought about two unexpected consequences: fixing the public universities' ranking and causing a rise in commercial schools coaching for entrance examinations. Now that one common test measured all public universities/colleges applicants, these schools could be ranked according to the average scores of their applicants. The ranking was fixed, and repeatedly reinforced by the results of the Common Test every year. Another unexpected but notable consequence of the Common Test was the rise and proliferation of coaching schools (YOBIKO). Because the applicants themselves had to decide which university to apply to, based on the self-scored results of the test, they needed to know the mean score of all the applicants, as well as the applicants of the universities of their choices of the given year they took the test. That is, although the Common Test was intended to serve as a criterion-referenced test measuring the core knowledge of the high school English syllabi, it was actually used for selective purposes, and an applicant's being accepted by a particular university depended on the scores of the applicants of that particular year. The coaching schools that had mock test results given at their nation-wide network of schools for the previous year, were the only ones that could provide an approximation of such expected mean scores of the applicants who were most likely to take the Common Test. The applicants' desire to know which universities were likely to accept them on the basis of the expected applicants' mean scores was very strong because such information made them better able to prepare for
the second entrance examinations that varied from university to university. Because the quality of the expected mean score information provided by the coaching schools tended to be quite high, high school students as well as their teachers, started to rely heavily on such information, which led to a steep growth of the coaching school industry during Period 3.

In contrast to these rather negative consequences of Japan's economic growth, positive effects to English classrooms also occurred in Period 3. In 1964, overseas travel became liberalized for the first time since the end of World War II, but it was not until the beginning of Period 3 that Japan started to observe a massive increase in overseas travel. In 1964, only 127,749 Japanese traveled abroad, whereas in 1972, as many as 1,392,045 did, which was 430,000 people more than the previous year's overseas travelers (in 2003, 16,500,000 people went abroad). Some of them stayed for long periods, and some worked abroad. During such stays, Japanese for the first time saw English actually spoken and used for communicative purposes (Suzuki & Wakabayashi, 1999). This phenomenon occurred together with the internationalization of English. As Crystal (1997) and McKay (2002) have pointed out, English has won the status of "the international language par excellence (italics in the original)" (McKay, 2002, p. 5) especially after the World War II, and its power has become even stronger after the 1980's, due to the accelerated globalization of business, culture, and information. It was thus a natural consequence that people started to doubt the content validity of the traditional grammar-translation teaching method that had continued to be dominant in English classrooms in Period 3.

With such public sentiment as a background, in April, 1974, Mr. Wataru Hiraizumi, the vice chair of the ruling Liberal Democratic Party's Special Committee on International Cultural Exchange submitted "The present situation of teaching foreign
In the plan, Mr. Hiraizumi criticized the current state of foreign language education in Japan as "useless" because "on the day after their graduation, (senior high school) students can hardly read, write, or understand the foreign language they have studied (for six years)" (Hiraizumi, 1974, p. 10; translated by M. Sasaki). He further argued that studying such useless English had now become "a necessary evil" because almost all senior high schools and universities required it as a subject of their entrance examinations. Among the steps he proposed for improving such a situation in the Plan were (1) to abolish English as a junior high school subject, and replace it with a subject such as "Languages and Cultures of the world;" (2) to make the senior high school English classes elective; (3) to select 5% of the high school applicants for English education, and make them study English intensively for professional purposes; and (4) to remove English from the entrance examination subjects. Since then, Mr. Hiraizumi's plan has had a profound and long-lasting impact on English language education in Japan. For example, The English Teachers' Magazine, one of the most popular English education magazines in Japan, featured such issue as "Useful English Education" several times after Hiraizumi's plan was published (e.g., the April, 1975 and August, 1978 issues). Quite a number of English teachers also presented various counterarguments, such as the one presented by Professor Shoichi Watanabe at Sophia University, maintaining that foreign language education is intended to cultivate students' mind by forcing them to translate the original foreign language texts into their mother tongue, and vice versa (see Hiraizumi & Watanabe, 1995), which had traditionally been believed to be the major purpose of EFL education in Japan (recall Period 1).
Few of Mr. Hiraizumi's suggested plans were fully realized today, but it did influence the government's policy towards English education in Japan because his plan reflected the general public's dissatisfaction with the lack of clear results for six years of high school English education in Japan. Responding to such negative sentiment, the Ministry of Education implemented several measures in Period 3. It introduced a new subject titled "English conversation" in the Course of Study for senior high schools promulgated in 1970. In the next Courses of Study promulgated in 1977 for junior high schools, and in 1978 for senior high schools, the contents to be taught in English classroom became further simplified and more practical. In terms of assessment, too, many more senior high school entrance examinations started to include listening comprehension items in the 1970's (e.g., in 1978, 26 out of 46 prefectures made tests with listening items, and in 2006 all prefecture tests had listening items). Furthermore, motivated by an idea similar to Mr. Hiraizumi's (1974) plan, an intensive English courses (Eigo-ka) were newly established at some publish senior high schools in 1977. Finally, towards the end of the 1970's, Wilkins' (1976) "notional syllabus" and "communicative language teaching" based on that syllabus were introduced (e.g., a series of special issues of *English Teachers' Magazine* featuring "Practice of Communicative Teaching" was published in 1978), and notion/situation-based texts began to be used for both teaching and assessment materials especially in 1980's.

The last effect of Japan's rapid economic growth on English education in Period 3 was the government's hiring of large numbers of English-speaking teachers. This was mainly conducted through the JET (Japan Exchange and Teaching) Program, which started in 1987 originally to spend part of Japan's trade surplus. In the first year of the JET Program in 1987, the government hired 848 English-speaking teachers from countries such as Australia, New Zealand, the United States, and England. These
teachers are called ALTs (Assistant English teachers), and as of 2002, there are a total of about 8400 ALTs, including 2794 hired by local municipal governments, and not by the JET program (Kan, 2002). Such English-speaking teachers are usually stationed either at junior/senior high schools, or at communities containing several high schools, and visit English classes regularly. They are usually asked to team-teach with the full-time Japanese English teachers at the given schools. Although there have been problems, such as the difficulty of selecting well-qualified teachers (e.g., Kan, 2002), the introduction of English-speaking teachers to the classrooms, alongside the popularization of overseas trips mentioned above, have increased people's motivation to learn (especially spoken) English for "practical" purposes.

Toward the end of Period 3, education related problems such as juvenile delinquency, bullying, school absence and dropout became serious (The Ministry of Education, 1989). The Ministry of Education tried to solve these problems by their educational reforms exemplified in the new Courses of Studies promulgated for junior and senior high schools at the end of Period 3 in 1989, which led to Period 4.

PERIOD 4 (1990 to the Present): INTRODUCTION OF INNOVATIVE POLICIES

Japanese education in Period 4 can be best characterized by the introduction of the government's new educational policies and public criticism of the results of these policies. Although this period is shorter than the others, it has certainly been a period of "storm and stress." During Period 4, two sets (the 1989 and 1998 versions) of Courses of Study promulgated by the Ministry of Education (the Ministry of Education, Culture, Sports, Science and Technology, from January 6th, 2001; henceforth the MEXT, its official abbreviation) were put into effect, but they shared similar principles, often summarized with the three key phrases, "redefinition of academic ability," "introduction
of criterion-referenced evaluation system," and "further advancement of liberal, flexible
and comfortable school life (YUTORI)" across all subjects.

First, the academic ability to be achieved during the elementary and secondary
education was redefined as "motivation/attitude for learning and the ability to solve
problems as a autonomous person in response to societal changes (Kariya, 2002, p. 56;
translated by M. Sasaki)." This ability has also been summarized as "zest for living"
(the Ministry of Education, Science, Sports and Culture, 1999). Such a redefinition
was a reaction against the past emphasis on the cramming of facts, especially for
entrance examinations, and has been interpreted by educational researchers (e.g.,
Mizukoshi, 1993) as a "paradigm shift" from the traditional view treating academic
ability in term of its substance (e.g., knowledge and skills) to the one treating it in terms
of its function (Abiko, 1996). Accordingly, the purposes of English (or foreign
language in general) education written in the Courses of Study for 1993 and 1994
included cultivation of "the attitude toward active involvement in communication"
(translated by M. Sasaki), in addition to the original purposes of cultivating
understanding of language and culture, and language skills themselves.

Furthermore, the new Courses of Study required teachers to evaluate all abilities
and knowledge, including "attitude," according to a criterion-referenced measurement
system. The criterion-referenced system was newly advocated to make sure that all
students have learned "basics and fundamentals with sense of achievement" (the MEXT,
2002). At the actual teaching level, the School Guidance Record revised for the 1989
version of the Course of Study required junior (but not senior) high school English
teachers to give an A (Satisfactory), B (Fairly satisfactory), or C (Need effort) according
to the following four criteria (translated by M. Sasaki):
(1) Interest, motivation, and attitude toward communication: Interested in communication, and willing to communicate.
(2) Expression: Speak and write one's opinions using beginning English.
(3) Comprehension: Understand the speakers' and the writers' intention by listening and reading beginning English.
(4) Knowledge and understanding about language and culture: Understand the language and thought and culture behind the expressions through learning beginning English.

In addition, the teachers had to give scores of 1 to 5 for students' overall English proficiency by means of a "norm-referenced system with some flavor of a criterion-referenced system (with a warning that teachers should be careful not to follow the conventional normal curve)” (translated by M. Sasaki). In the next version of the Course of Study that is still in effect now, in addition to giving an A, B, or C according to the criterion description of the four criteria similar to those above, teachers are required to use a complete criterion-referenced system for their students' overall English proficiency, from 1 (Needs much effort) to 5 (Fully satisfactory).

Because the teachers who actually had to evaluate students had been accustomed to norm-referenced measurement systems (ever since Period 1), but not to criterion-referenced measurement system, this “Copernican change” (Matsuzawa, 2002, p. 40) resulted in a great confusion. Even in the year 2002, more than 10 years after this system was first introduced, there were complaints such as "The evaluation criteria are not transparent," "Teachers vary greatly in their evaluations," and "The given scores are not useful for senior high entrance examinations" (Matsuura, 2002, p. 10; Translated by M. Sasaki). It is easy to imagine how difficult it would be to evaluate students according to criteria such as "interest, motivation, and attitude" (Wakabayashi, 1993). Furthermore, it may put an additional burden on the teachers if they must evaluate students using eight sub-criteria (of the four main criteria presented above) more than twice, as Matsuura suggested for 40 students every semester (high school teachers usually teach four to five classes). As Negishi (2005) pointed out, many problems still
need to be dealt with (e.g., establishment of shared understanding of all criteria
descriptions, and establishment of shared criteria across schools) before this system can
be successfully employed in all high schools in Japan.

The third key phrase characterizing the MEXT's new educational policies in Period 4 was "the further advancement of liberal, flexible, and comfortable school life." This was most notably reflected in the 1998 versions of the Courses of Study (effective in 2002 for junior high and in 2003 for senior high schools), where the complete five-day school week and a 30% cut in curriculum content took effect. A content reduction has been suggested ever since the 1958 version of the Course of Study, but this most recent one was the most drastic. Furthermore, as part of the measure taken for this policy, the Ministry of Education in 1993 decided to ban the use of commercial tests and standardized scores called HENSACHI (with the mean of 50 and the SD of 10) at junior high schools. This was intended to alleviate the "excessive competition in entrance examinations" (Ministry of Education, 1989). These reform measures were intended to achieve "liberal, flexible, and comfortable school life," however, have not worked as the Ministry originally intended. For example, banning the use of commercial tests and HENSACHI only made junior high school students and teachers depend all the more on private coaching schools. The Ministry did not realize that the information provided by those coaching school had by then become a kind of a "necessary evil." Based on the nation-wide branch-school networks, the coaching schools, especially the big three (Kawaijuku, Yoyogi, and Surugadai), have been the only existing institutions in Japan that monopolized reliable information such as the predicted cut-off of senior high schools and universities of the applicants' choice for guiding entrance examination applicants. This policy has resulted in an increasing number of students who go to
coaching schools after their regular schools, and the number of those coaching schools themselves.

The MEXT's decision to reduce the curriculum content by 30% has also generated adverse effects. Researchers such as Kariya (1995, 2002, 2003) and Okabe, Nishimura, and Toze (1999), using concrete examples, have pointed out that the changes in the Japanese education since the introduction of the concept of the "liberal, flexible, and comfortable school life" have resulted in a serious deterioration in children's academic ability. The results of the MEXT's own study conducted in 2001, provided additional evidence that the academic ability of children in compulsory education (ages from 6 to 15) decreased in many subjects such as social studies, mathematics, and science between 1993 and 2001. In addition to the deterioration of the overall academic ability, researchers (e.g., Editorial Board of *Chuoukouron* and Nakai, 2001; Kariya, 2002) have also reported that there was a decrease in children's study time after the "liberal, flexible, and comfortable school life" policy took effect at the end of the seventies. Even more seriously, Kariya (1995; 2002) discovered that such a decrease in children's study-time reflected the socio-economic status of their parents, and that the study time of children in lower economic levels tended to drop most steeply. He warned that the MEXT's "liberal, flexible, and comfortable school life" policy has increased the gap between the rich and the poor, which is likely to create future social unrest.

Reacting to these negative aspects and the public criticism, the MEXT decided to review their educational policy in 2000, and it has subsequently announced a series of new reform plans. For example, on October 11th, 2000, Education Minister Mori commented on the Sankei newspaper that the curriculum content written in the Course of Studies is the "minimum standard," and that it can be expanded when taught in classrooms. Considering that it had been interpreted as the "maximum standard" by
many school teachers until then, this was quite a notable interpretation (Kariya, 2003). After that, the "Educational Reform Plan for the 21st Century" was announced in 2001, the "Human Resources Strategy Vision" in 2001, and the "Compulsory Education Reform Plan" and the "Japan! Rise Again! Plan" in 2004. Having inherited the fundamentals from the previous policies, those new reform plans clearly show a directional reorientation toward meritocracy, emphasizing "rivalry among students" (Japan! Rise Again! Plan), and the "fostering of top-level talents who will lead the Century of Knowledge" (Human Resources Strategy Vision) aspects, which the Ministry of Education once seemed to have tried to eliminate from the classrooms.

Finally, in addition to influence from politically instigated educational reforms, the continuing progress of globalization has also influenced EFL education and assessment in Japan in Period 4. Because of the accelerating speed of globalization, the MEXT has maintained that “children need to acquire communicative competence in English, which is an international language, in order to survive in the 21st century,” but also that “many Japanese refrain from communicating with foreigners, or underestimated because of their poor English ability” (translated by M. Sasaki; the MEXT, 2002). Based on such belief and opinion surveys, the MEXT announced the "Strategic Plans to Cultivate Japanese who Can Use English” (translated by M. Sasaki) in 2002, and has consequently put some of their “Action Plans” into practice. These plans include the introduction of a listening subtest to the Center Examination mentioned in the beginning of this paper, further emphasis on practical communication abilities, building 100 “Super English High Schools” over three years, the enrichment of English conversation activities at elementary schools, and requiring TOEFL (Test of English as a Foreign Language) scores of 550 or TOEIC (Test of English for International Communication) scores of 750 for all English teachers.
The idea of emphasizing practical (especially spoken) communicative abilities in EFL classrooms was already emphasized in Period 3, but the MEXT policy in Period 4 has added further prominence to this idea in the two Courses of Study promulgated in Period 4. Thus, in the 1989 versions, the new subjects entitled “Oral Communication A, B, and C” were introduced to be taught with other subjects such as “Reading” and “Writing,” and grammatical items became weighted less for high school students. Furthermore, in the 1998 versions that are still effective, the phrase “practical communication ability” was added to the general purpose of the English subject for high school students. The notional-functional aspects of language use (termed "language-use situations," such as traveling and debating, and "language functions" such as greeting and thanking) are also highlighted with quite a number of concrete examples in these Courses of Studies. Such changes in the content to be taught at high schools have also affected the content of English language assessment although the changes are slow. As we have seen in the "Strategic Plans to Cultivate Japanese Who Can Use English," commercially-based tests measuring some type of "practical" English abilities, such as TOEFL, TOEIC, and STEP (The Society for Testing English Proficiency established in 1963) tests, have now been recommended to replace part of senior high school and university entrance examinations (the MEXT, 1999). As for the university entrance examinations themselves, Negishi (1997) reported changes, such as an increase in conversation texts (instead of literary texts) as test materials, free compositions (instead of Japanese-English translation), gist-searching items based on relatively long texts (500 to 1500 words) as possible effects of the 1998 Course of Study. Accordingly, what is done in the English classrooms preparing the students to take these examinations is also expected to change.
Finally, the rapid decrease in the number of school-age children in Japan, especially high school graduates, has also influenced English language assessment in Period 4. After the post-war baby boomers’ (about 8 million people born between 1947 and 1949) children were born in the sixties and seventies, the birthrate has continued to fall (it was 1.29 in 2003). Meanwhile, the number of universities/colleges has continued to grow (e.g., 460 universities in 1984 to 622 in 1999). Consequently, the competition for entering higher education has become much less severe in this period, and 60-70% of senior high graduates have proceeded to post-secondary education (e.g., in 2003, 41.3% went to universities, 7.7 to junior colleges, 23.1 to vocational colleges). The university passing rate (the ratio of those who passed against those who took the examinations) rose to 80.5% in 2000, from 55.5% in 1990 (it was 42.1% in 1959). Furthermore, since the National Council on Educational Reform recommended "diversification of selection procedures for upper secondary school applicants" in 1984, almost all universities/colleges have started to employ procedures (e.g., interviews, short essays, recommendation by high school principals, or admission office examinations that can include these or other evidence) other than high-stake one-time tests. Many universities have also decreased the number of subjects required for entrance examinations. These steps have greatly eased the severity of preparing for upper secondary entrance examinations, but have also lowered the applicant's motivation to study. Kariya (2002) reported that senior high school students hoping to enter universities through their principals' recommendation (without taking the one-time examinations) tended to study much less than those who plan to take the regular high-stake examinations. It is also possible that such study-time decrease may have caused the past deterioration of Japanese students' academic ability as well.

SUMMARY AND IMPLICATIONS
Figure 1 summarizes the 150-year history of EFL assessment in Japan described above. In Period 1, the global social/political pressure forced Japan's to open its closed doors to the world, and Japanese students started to study English at school for the first time in Japan's history. However, it took only 20 years before nationalism prevented Japanese from studying English as a means of communication. After English became a school subject studied exclusively by social elites, teachers began to adopt grammar-translation methods, which had long been successfully used for teaching/learning Chinese as a written language. In Period 2, the most influential global socio-political factor was Japan's defeat in the World War II, which consequently made EFL education essentially compulsory after the war. This resulted in many unqualified EFL teachers who were forced to depend on the old grammar-translation methods that they had inherited from the prewar period. Overall, EFL education and assessment up to the end of Period 2 tended to be passive in terms of content and method, trying to one-sidedly absorb foreign knowledge and culture through English. In contrast, in Period 3, the Japanese people's purpose of learning English became more bilateral/communicative because of their economic prosperity as well as internationalization of English. The content of EFL teaching and assessment became more practical and speech-oriented. Meanwhile, the increase in family income resulted in an increase in the percent of those who proceeded to higher education, which in turn produced problems such as school dropouts and excessive competition. Responding to such problems, the government in Period 4 tried several innovative reforms, including the promotion of multiple assessment methods and criterion-referenced measurement systems. In this period, internationalization of English in socio-economic globalization has further accelerated, which has made EFL education and assessment in Japan even more practical and communicative-oriented. To sum up, the four periods listed in Figure 1 were distinct
from each other in terms of the "intended goals and the degree of popularization of school-based English education in Japan." Moreover, as I briefly recapitulated above, the school-based English language assessment practices in these periods have been influenced in different ways by the socio-political, economic, academic, and historical-cultural factors. These findings confirm my two initial working hypotheses.

In Figure 1, all factors, stronger or weaker, depending on each period, have arrows directing toward "EFL assessment." However, as many researchers (e.g., Lynch, 2001; McNamara, 2001, see also Watanabe, this issue) have recently argued, the arrows can also point in the opposite directions. In fact, many arrows in Figure 1 can have bi-directional pointers with the other factors. One example from the present study will suffice. While the Japanese government's decision to give the Common Test in Period 3 had a significant impact on teaching and assessment in senior high school EFL classrooms, at the same time, without it being the government's intention, it also greatly benefited private coaching schools (an organization belonging to the Economic Factor) that had monopolized information about the applicants' expected scores for the Common Test. Furthermore, because all coaching schools charge fees for providing such information, this may have created domestic social/political problems by unfairly providing such information only to those who can afford it.

Such complex relationships between EFL assessment and other factors reminds me of Atkinson's (2002) metaphor of the environment surrounding language learners, as "a tropical forest, so densely packed and thick with underbrush that it would be hard to move through" (p. 526). He used this metaphor for fields such as "language socialization and cultural anthropology" (p. 525), but it also seems well fit for the environment surrounding EFL assessment described in this paper. Although historical studies are rare in the field of language testing, this similarity suggests potentially
promising areas to be studied in the future. Just as Atkinson suggested that a phenomenon associated with second language acquisition can be studied as a complex "sociocognitive" (p. 524) entity as a whole, we can expect that a phenomenon associated with language assessment can also be studied as a complex entity as a whole. For example, we could investigate why the introduction of a criterion-referenced measurement system into EFL high school classrooms in 1993 created such confusion. As a framework, we could use Fullan's (1991) innovation theory advocated by Wall (1996) by considering "the three R's of relevance, readiness, and resources" (cited in Wall, p. 339). Briefly recalling the recent history of EFL assessment in Japan suggest several potentially related facts such as the fact that many teachers did not think the new measurement system would be "relevant" to them because their students still had to take norm-referenced entrance examinations to get into senior high schools, or that the system was not fully ready when it started because there did not exist detailed descriptions for each criterion shared across different schools. On the basis of such analyses, we could suggest useful measures to improve the situation, or we could approach the same phenomenon from a more critical perspective considering principles such as "the need to monitor the uses of tests as instrument of power, to challenge their assumptions and to examine their consequences" (Shohamy, 2001, p. 376). In fact, Kariya (1995) pointed out that complicated measurement systems tend to work unfavorably for those who are poor. The results of such historical studies are thus likely to provide information that traditional mainstream cognitive-oriented language testing research may have overlooked. (8860 words)
Author's Notes

This paper was presented at the 2006 International Conference on English Instruction and Assessment, National Chug Cheng University, Taiwan, on April 23rd, 2006. I would like to thank Bill Herlofsky and Paul K. Matsuda for their valuable comments and suggestions. Needless to say, I am solely responsible for the positions taken and any errors that may remain.
1 In spite of such intensive preparation, the Center had to retest 436 examinees (.09%) because of the tape player troubles, which was publicly criticized for its "high" trouble rate.

2 For example, in 1866, one year before the end of the Tokugawa Period, it admitted 150 English-major students.

3 In 1874, the school became part of the Imperial University of Tokyo, the only university in Japan then.

4 An example of an accent location item follows:

Choose the word with the strongest accent for the underlined part:

(a) an-oth-er (b) chil-dren (c) Sat-ur-day (d) some-thing

An example of a written pronunciation item follows:

Choose the word that has the same pronunciation as the one underlined in the given word:

said (a) came (b) clean (c) may (d) breakfast

5 This new policy was met by fierce criticism from English teachers for lowering the students' English ability (see The English Teachers’ Magazine, 1981, March and October issues, for example). It has made private junior high schools more popular where many more hours of English instruction are given, which has thus widen the gap between the rich and the poor in terms of receiving a good education (Asano, 1986)

6 After 1990, when its name became the Center Examination, private universities started to adopt the examination, and in 2006 a total of 303 universities, about a half of all universities in Japan, used the examination.
7 However, the entrance examinations for private senior high schools have been long criticized by their difficult content that is beyond the scope of the curriculum content suggested by the Course of Study (see The English Teachers’ Magazine, 1984 April issue, for example).

8 In this study, the first and second junior high students significantly improved their English scores, but the freshman scores became significantly worse.
150-Year History of School-Based EFL Assessment in Japan /p. 35
References


Editorial Board of Chuoukouron and Nakai, K. 2001: *Ronsou: Gakuryokuhoukai [Controversies over the deterioration of students' academic ability]. Tokyo: Chuoukouronshinsha.*


150-Year History of School-Based EFL Assessment in Japan /p. 37


Hayama, 1963:


150-Year History of School-Based EFL Assessment in Japan /p. 38


Kan, M. 2002, November: ALT ga fueru nowa iikeredo: Koe ni dashite yondewa ikenai ALT mondai [Although it is good ALT to increase...: ALT problems you should not read aloud]. *The English Teachers' Magazine*, 16 - 17.


------. 2003: *Naze kyouikuronsou wa fumou nanoka: Gakuryoku ronsou o koete* [Why are controversies over education so unproductive? Beyond the controversies over academic ability]. Tokyo: Chuoukouronshinsha.


---. 1999: *Japanese government policies in education, science, sports and culture.*


------. 1997, September: Shin karikyurumka deno daigakunyuushi: Dou kawattaka, dou kawaruka [University entrance Examinations under the new curriculum: How have they changed, and how will they change?]. *The English Teachers' Magazine, Special Issue*: 4 - 9.


Saito, E. 2002, November: Sentaa shiken ni iyoioy lisuningu donyuu ka [Is the Center Examination finally going to include a listening subsection?]. *The English Teachers' Magazine*: 22 -23.


Watanabe, Y. This volume.


150-Year History of School-Based EFL Assessment in Japan

- **EFL Teaching/Assessment Period 1**: EFL for Elites (1860-1944)
- **EFL Teaching/Assessment Period 2**: EFL for Everyone (1945-1969)
- **EFL Teaching/Assessment Period 3**: EFL for Communication (1970-1989)
- **EFL Teaching/Assessment Period 4**: EFL under Educational Reforms (1990-Present)
Figure 1. Factors affecting the EFL teaching and assessment in Japan across the four periods