The Economic Journal of Takasaki City University of Economics vol.49 No.1 2006 pp.43～63

Improving Transcription for Japanese Learners of English

Yuzawa Nobuo

Keywords: English, transcription, pronunciation, IPA, IPA-J

0. Introduction

Some Japanese claim that English pronunciation is not important. If they mean by this that English pronunciation with a Japanese accent is acceptable as long as it is communicable, this is an appropriate statement.[1] If they mean, however, that careless, sloppy pronunciation is acceptable, they are wrong. Bad pronunciation makes it difficult or even impossible to communicate successfully. Therefore, it is important in English teaching to enable Japanese learners to pronounce English like native speakers or at least communicably, and currently great effort is being made in several educational sectors across Japan to achieve this purpose.

Recently there has been a movement of implementing English teaching in Japanese public elementary schools[2]. The author thinks that it is desirable to input English sounds into young children's brains before they reach the 'critical period'. It seems that humans learn to perform different tasks in different developmental stages and that each learning capability degenerates after passing a certain period of time. For example, a child who was brought up by a wolf was unable to acquire a human language, and a person cannot acquire 'perfect pitch' unless he/she takes music lessons when he/she is small. If English teaching in elementary school ends in success, many problems with English pronunciation (e.g. the distinction between /l/ and /r/) may be solved entirely.

It is unfortunate that the present English teaching system in Japan is not as advanced as this, but teachers need to deal with their current students and do not have time to wait for the ideal teaching method to be realized. They have to be constantly looking for ways to improve. They need a better method of teaching English pronunciation to Japanese learners of English, especially those who have passed the critical period. This paper deals with effective ways of improving such learners' English pronunciation, particularly its segments, from the viewpoint of transcriptions.

1. SET theory

The author thinks that in order to enable Japanese learners of English to acquire English pronunciation successfully, three elements need to be integrated effectively: (1) 'sound' as stimuli, (2) 'transcription' to make those sounds easier to acquire, and (3) 'evaluation' to judge learners' pronunciation accurately. This integrated English pronunciation teaching method, called 'SET theory' ('S' for sound, 'E' for evaluation and 'T' for transcription), is diagramed in Fig. 1:

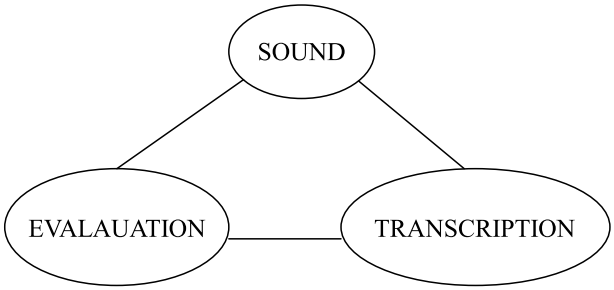

Fig. 1 SET theory

In the following, the three elements are explained.

2. Sound

It goes without saying that among the three elements, sound is the most important. The other two simply work complementarily to make pronunciation teaching successful. In real communication, neither transcription nor evaluation is necessary, and the role of transcription can only be found in checking how to pronounce unknown words in dictionaries. Without these two elements, however, it is difficult to acquire English pronunciation in non-English-speaking countries like Japan.

An important question that first comes to mind is which English pronunciation should be used as a model. To find a key to answering this question, the author refers to what is written in ***The Course of Study*** because Japan's educational policy is delineated there. The Ministry of Education (1999a: 33) states the model should reflect 'today's standard pronunciation' and continues to say that 'Today, English is widely used in the world, but it is used differently in various settings and many varieties exist in its pronunciation and usage. What should be taught among today's various

English pronunciations is the so-called 'standard pronunciation' which is not limited to specific areas and groups and is not too colloquial (the author's translation and emphasis).' But this explanation does not make sense. ***The Course of Study*** prohibits choosing English pronunciation from specific areas and groups, but it is practically impossible to teach English pronunciation without targeting a specific area as a model. The Ministry of Education (1999b: 113) also mentions 'today's standard pronunciation' and there is no explanation of what that means.

Judging from this state of affairs, it seems that it is school textbook writers and editors that make a final judgment on which type of English pronunciation should be used as a model. Major English-speaking countries are the UK, the USA and Australia, and standard pronunciation used in those areas could be a model. In Japan's English school textbooks, GA is heavily emphasized. This reminds the author of an episode that he heard from a textbook writer. He wanted to use British English spellings when selected resource data was written in British English, but he was advised by the editors to change them to American spellings. An English listening comprehension test was first administered in 2006 in the Center Exam, but only GA was used. Thus, Japan's English teaching is American English centered at least up to university entrance exams, and 'varieties' mentioned in ***The Course of Study*** are neglected.

In spite of this lopsided approach, these 'varieties' are observed in selecting ALTs (i.e. assistant language teachers). They are selected from among various English-speaking countries. However, this creates another problem, because English spoken by an ALT in the classroom may be very different from the English coming from audio devices accompanying a school textbook. The situation is very serious when learners are junior high school students who are learning English for the first time. They may be perplexed as to which English pronunciation they ought to learn, and their teachers may also have trouble explaining the differences in pronunciation. This situation is far from satisfactory. The author has a similar experience. When he was team-teaching with a British teacher, one student in the classroom told him that the teacher's pronunciation was strange. His pronunciation was RP and this was the first time for the student to hear it. If this current system continues, these problems may be prevalent throughout Japan.

The author thinks that both GA and RP should be taught in school as they are the two major types of English pronunciation in the world and both are transcribed in English-Japanese dictionaries. However, this problem of selection will not be discussed any further here, because this is not the main theme of this paper. Judging from the reality of Japan's English teaching, the model pronunciation will be regarded as GA in this paper.

3. Evaluation

Evaluation is included in the SET theory. It is important to judge accurately whether each learner's pronunciation is correct and give him/her appropriate feedback. Ideally, teachers should monitor each learner's pronunciation during the class and give pronunciation tests regularly, but it is practically impossible to make such painstaking efforts with 40 or more students in a classroom.

To improve this situation, an automatic pronunciation evaluation system may be an option if it is financially feasible and technology continues to advance. A few years ago, the author bought a PC software program for pronunciation evaluation, but it was not good enough. He intentionally inserted a vowel at the end of a word-final consonant[3], and pronounced the word. He passed the test. This kind of error detection program may not have been installed in the software, but programs to detect errors commonly found in Japanese learners must be installed. The author has not bought another software program of this kind since then, but he hopes that accuracy level has been greatly improved.

The author supports the use of an automatic pronunciation evaluation system if financial and technological problems are solved. There are three reasons for this. First, learners can practice English pronunciation anywhere anytime. Second, those who do not like to practice in front of others can be relieved of psychological stress. Third, learners can get appropriate feedback to improve their pronunciation.

4. Transcription

The third and main element in the SET theory is transcription. There are three approaches: the quantitative approach, the qualitative approach, and the quantitative-qualitative approach. To take 'pool' and 'pull' — a minimal pair including the close back vowels — as examples, they are transcribed as /ˈpuːl, ˈpul/, /ˈpul, ˈpʊl/, /ˈpuːl, ˈpʊl/, respectively. What should be noticed here is that /ˈpul/ means two different words in different approaches: 'pull' in the quantitative approach and 'pool' in the qualitative approach. This misunderstanding may not be a problem because the qualitative approach is not common in Japanese English teaching or English-Japanese dictionaries.

English-Japanese dictionaries adopted the quantitative approach in the past, but many of them have changed to the quantitative-qualitative approach[4]. This is a good shift, because vowel duration is not fixed in English and qualitative differences become more important. When a syllable has a structure of VC, the vowel is shortened when the consonant is voiceless or fortis.[5] To

represent this phonetic reality, it is obvious that the quantitative-qualitative approach is most ideal, but this approach is not used in school textbooks approved by the Ministry of Education[6].

Both the quantitative approach and the qualitative approach are theoretically correct. The point is, which is made explicit and which is made implicit. It is expected that in the quantitative approach, quality differences are inferred, while in the qualitative approach, quantity differences are inferred. Since Japan's school textbooks adopt the quantitative approach, it is imagined that the writers and editors assume that learners are already familiar with quality differences, but the author does not think this is the right assumption as far as his experience is concerned. In Japanese, vowel quantity is important in distinguishing, for example, 'ojiisan' (grandfather) from 'ojisan' (uncle). There is an accentual difference between the two words, but vowel quantity is more important than the accentual difference. Judging from the fact that there are only five vowels in standard Tokyo Japanese, it is easily presumed that Japanese learners of English in general are not sensitive to differences in English vowel quality. Needless to say, the quantitative-qualitative approach is the most ideal for Japanese learners of English.

However, the quantitative-qualitative approach is not problem-free — the method of transcription is not unified. Symbol use varies, depending on the English-Japanese dictionary, and some dictionaries use their own unique symbols. Each learner should be aware of these differences in advance because it is practically impossible for all Japanese publishing companies to use the same transcription system. By pointing out such problems, one by one, the author would like to discuss which transcription is better for Japanese learners of English.

The first problem is how to transcribe the vowel in 'turn'. In English-Japanese dictionaries, either /ə:r/ or /ɚ:/ is used, but they may give different impressions to learners. The former is likely to mean pronouncing the schwa strongly and then /r/, which is a wrong interpretation. In contrast, the latter is likely to mean pronouncing the /r/-colored schwa strongly. Since the vowel in 'turn' has the same quality from beginning to end, the latter transcription is easier to understand. Phonemic transcription is used in dictionaries, but the author wonders if learners can easily understand the phonetic realization of /ə:r/ as [ɚ:]. It is also unknown whether learners can understand that /ə:r/ is totally different from /ɑ:r/ or /ɔ:r/, where there is a gradual change in quality from the vowel to /r/. Theoretically, it is important to represent a system as economically as possible, but it is also important for dictionary writers and editors to incorporate necessary phonetic facts and to make dictionaries as learner-friendly as possible. It is desirable to introduce /ɚ:/ as a phonemic symbol for Japanese learners of English.

The second problem is /ɚ:/ itself. In the above, the author talked about the validity of this symbol, but it is not faultless. Learners may think that this has something to do with the schwa.

This is right in terms of acoustic image, but the strong/weak distinction in syllable and vowel is important in English. Neither the schwa nor the /r/-colored schwa ever appears as a strong vowel in a strong syllable. The author wonders if learners can understand this important difference with only the length mark. Roach et al. (2003), Upton et al. (2001) and Wells (2000) transcribe this vowel as /ɜ˞ː/, and the author is also in favor of this symbol. So far, this has never been used in English-Japanese dictionaries, but it should be introduced as soon as possible. This symbol clearly shows that not only is its vowel quality the same from beginning to end but also it is never a weak vowel. Some may claim that this symbol may confuse learners, but the author believes that the addition of only one symbol is not a big problem. On the contrary, if it is used in English-Japanese dictionaries, it will be advantageous for learners when they begin to use English dictionaries. All the symbols used there would be already familiar to them.

The third problem is how to transcribe the vowel in the second syllable of 'mother'. Basically, there are two types of transcription of this vowel: /ər/ and /ə˞/. In the former, it is expected that learners are aware of the meaning of /ər/ as [ə˞], just like /əːr/ as [ə˞ː]. Not all learners, however, may be able to understand this phonetic realization. The author recommends that /ə˞/ should be used instead. This symbol is also adopted in Roach et al. (2003).

The fourth problem is italic /r/. This symbol was invented in Japan to transcribe GA and RP together. This is helpful in economizing available space in dictionaries, but its use should be stopped because learners may misunderstand this symbol (Takebayashi (1996: 294-295)). There are seven symbols transcribed with this italic /r/ and they mean as follows: (1) /əːr/ = /əːr | əː/, (2) /ɑːr/ = /ɑːr | ɑː/, (3) /ɔːr/ = /ɔːr | ɔː/, (4) /ɪər/ = /ɪər | ɪə/, (5) /eər/ = /eər | eə/, (6) /ʊər/ = /ʊər | ʊə/, and (7) /ər/ = /ər | ə/[7]. GA is shown on the left and RP on the right. The author thinks that it is much easier to understand what each symbol means when they are transcribed separately rather than when they are transcribed together with the italic /r/. This complicated rule not only confuses learners but also is harmful to them. Takebayashi (1996: 295) states that he found an introductory textbook for English phonetics, where /əːr/ is used only to show GA. This is a good example of showing how serious the problem of the italic /r/ is. Learners who use this kind of textbook will never learn what GA actually is.

The fifth problem is how to transcribe /r/-colored vowels in 'fear', 'fair' and 'poor'. In some English-Japanese dictionaries, these vowels are regarded as diphthongs, by emphasizing phonetic diphthongization, and the second element is transcribed as /ə˞/. As Cruttenden (2001: 85) and Ladefoged (2001: 28) explain, however, these vowels should be transcribed as a short vowel + /r/. This matches the native speakers' mental image. This image is important in deciding phonemes. To take the consonant part of the Japanese /si/ as an example, this is phonetically different from the other consonants in the /sa/ column, but ordinary Japanese are not conscious of this difference.

The palatalization of /s/ before /i/, therefore, is not relevant in categorizing this consonant as a separate phoneme. In the same way, there is no logical reason why these /r/-colored vowels should be regarded separately. If this treatment is adopted, the number of GA phonemes is reduced and this helps to lessen learners' learning load.

5. IPA-J

The symbols used in dictionaries are phonemic symbols and predictable phonetic information is omitted. Dictionaries do not always describe learner-friendly information. There is no problem if learners can understand the meaning of symbols correctly. To take 'test' as an example, the two /t/s are realized differently, but this information is not represented in dictionaries. It is assumed here that learners are already familiar with this phonetic change, but this assumption ends in disappointment when learners cannot read transcription well enough. To help Japanese learners of English to read phonemic symbols in dictionaries, the author has independently invented what he believes to be a useful transcription system for those learners. This is a broad phonetic transcription and includes minimum phonetic information that is essential to Japanese learners of English. This transcription is named 'IPA-J', which means the IPA for the Japanese.

The idea of IPA-J came to the author's mind while he was preparing for his presentation at a JACET English lexicography symposium in March 1999. This symposium was held to examine whether kana transcription, especially Approximate Kana Transcription (henceforth, AKT) invented by a Japanese prominent English phonetician, Takashi Shimaoka, is useful to Japanese learners of English. The author was asked to attend this as a presenter and he constructed his argument to make complementary use of kana and the IPA. During this preparation, he thought of applying the idea of AKT to the IPA. In AKT, kana style is changed in its size and type for the purpose of helping Japanese learners to pronounce English better. The author thought that this visual effect could also be used for the IPA. Fig. 2 shows the relation between IPA-J and AKT.

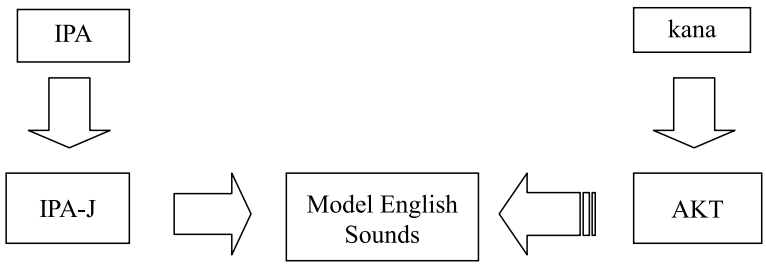

Fig. 2 IPA-J and AKT

The author's IPA-J (i.e. the visually modified IPA) and Shimaoka's AKT (i.e. visually modified kana) are merged here as a comprehensive transcriptional guide for helping Japanese learners of English to acquire model English sounds. These two methods of transcription function as if they were two wheels of a bicycle, but strictly speaking, the relationship is not a mirror image. IPA-J is targeted at model English sounds, but AKT aims at their approximation by showing direction to them. Fig. 3 illustrates this difference.

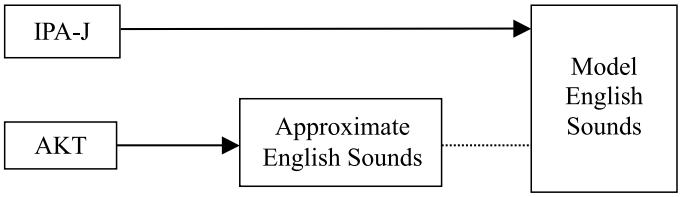

Fig. 3 Difference between IPA-J and AKT

This idea of aiming at the approximation of the model English sounds is related to interlanguage theory on which Shimaoka placed a theoretical background. To show this, part of one arrow is described differently in Fig. 2. IPA-J was also made public at the 4th conference of the English Phonetics and Transcription Association in 1999 and in Yuzawa (1999: 42-43). In Yuzawa (2000: 60-81), IPA-J was slightly revised and DST (i.e. Discrete Semitone) was also introduced to transcribe intonation. In 2006, IPA-J, slightly modified, was again made public at the 7th conference of the JACET English Lexicography Association. This is how IPA-J was invented and improved[8], along with its relationship with AKT.

There are five main features in IPA-J. The first feature is to make the first element of a diphthong more noticeable by displaying it larger than the second. There are no diphthongs in Japanese. When a vowel is followed by another vowel, they simply create a hiatus, not a diphthong. In English diphthongs, the first element is always stronger and longer, and this should be visually displayed. To illustrate this, spectrograms of Japanese 'ai' ('love') and English 'eye' are displayed in Fig. 4.

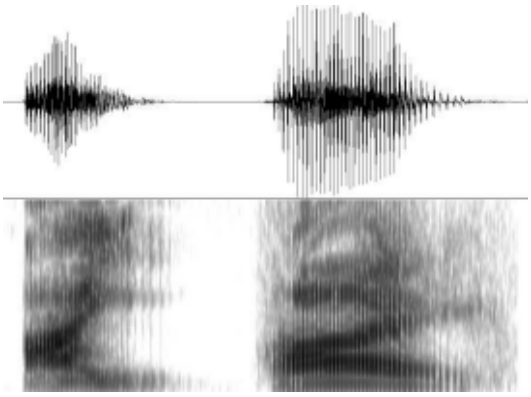

Fig. 4 Spectrograms of 'ai' and 'eye'

As this figure shows, difference in duration is noticeable between the two words. Japanese 'ai' (221 ms) is pronounced much shorter than English 'eye' (344 ms), but what is more important is the internal duration of these two words. In the case of 'ai', the first vowel (112 ms) is almost as long as the second (109 ms). In the case of 'eye', however, the first element (257 ms) is much longer than the second (87 ms). Japanese speakers perceive two units (i.e. two morae) in 'ai', while English speakers perceive only one unit (i.e. one syllable) in 'eye'. These two vowels are different in physical reality and mental recognition, and this difference must be made clear to Japanese learners of English. The GA diphthongs are transcribed in IPA-J as follows:

[aɪ, eɪ, ɔɪ, aʊ, oʊ]

The second feature is to make a strong syllable more noticeable by displaying it larger, in bold type and making it shaded[9]. These treatments are not applied to post-vocalic consonants in a strong syllable in order to visually avoid Japanese learners' bad habit of vowel insertion at the end of such consonants. Some GA examples are shown as follows:

'join' [**dʒɔ**ɪn][10] 'English' [**ɪ**ŋglɪʃ]

'drive' [**dra**ɪv] 'member' [**me**mbə˞]

'found' [**fa**ʊnd]

The third feature is to place stress marks before stressed syllables. This will help Japanese learners to capture the concept of syllables, as the basic unit of Japanese is morae. For example, 'strike' consists of one syllable, but this is pronounced in Japanese with five morae. 'MacDonald' is composed of three syllables, but this is pronounced in Japanese with six morae. Another advantage

of placing stress marks before stressed syllables is found in such examples as 'between', 'apply' and 'destroy'. If a stress mark is placed above the vowel in the stressed syllable, as in /bɪtwíːn, əpláɪ, dɪstrɔ́ɪ/, some Japanese learners may easily insert a vowel after /t/ in 'between', /p/ in 'apply' and /s/ and /t/ in 'destroy', and pronounce them as *[bɪtowíːn, əpulá ɪ, dɪsutorɔ́ɪ]. If such a stress mark is placed before the stressed syllable, as in /bɪˈtwiːn, əˈplaɪ, dɪˈstrɔɪ/, vowel insertion is less likely to occur and devoicing of a voiced consonant also becomes more understandable. Some claim that it is difficult to assign intervocalic consonants to a particular syllable. They may be assigned to the previous syllable or to the next. In this case, rules adopted in Roach et al. (2003) and Wells (2000) are useful. They adopt different rules, but the result is the same in these words. Roach et al. (2003) adopts the maximal onsets principle, which means that where possible, syllables should be divided in such a way that as many consonants as possible are assigned to the beginning of the syllable to the right, unless phonotactics is violated. Wells' principle is described in Wells (1990), which is that consonants are syllabified with the more strongly stressed of two flanking syllables. Some GA examples are shown as follows:

'between' [bɪˈtwiːn] 'apply' [əˈplaɪ]

'mistake' [mɪˈsteɪk] 'decision' [dɪˈsɪʒn][11]

'nonverbal' [ˌnɑːnˈvɜːbl][12]

The fourth feature is to display aspiration and the voiced /t/. Aspiration is important in English. Without enough aspiration, voiceless consonants may be recognized as voiced consonants. The voiced /t/ is a prominent feature in GA. To illustrate how important aspiration is, spectrograms of Japanese 'tai' ('versus') and English 'tie' are displayed in Fig. 5.

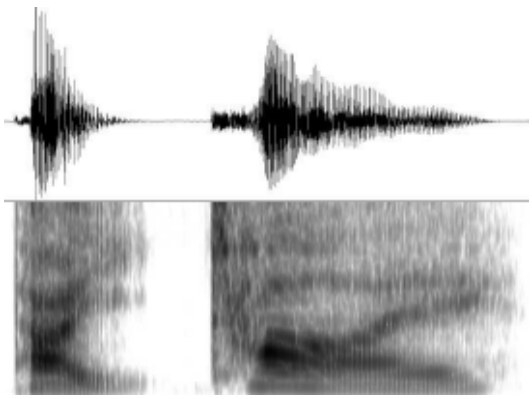

Fig. 5 Spectrograms of 'tai' and 'tie'

The left spectrogram shows Japanese 'tai' ('versus') and the right shows English 'tie'. There is a conspicuous quality difference in vowel: hiatus in Japanese vs. diphthong in English. As in Fig. 4, it is noticeable that the English word is spoken much longer than the Japanese word, the difference here being 622 ms vs. 256 ms. What is more important here, however, is the quantity difference in the prevocalic consonant: 99 ms in English vs. 33 ms in Japanese. In the duration of 99 ms in English 'tie', aspiration accounts for about 2/3 — 60 ms. In Japanese 'tai', however, such aspiration is nonexistent. This feature also plays an important role in semantic distinction in English (e.g. distinction between 'tie' and 'die'). Aspiration, therefore, deserves emphasis in learning English pronunciation.

The difference between the voiced /t/ and the normal /t/ is illustrated in Fig. 6.

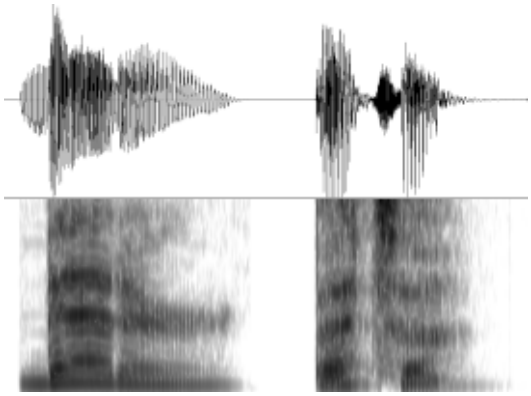

Fig. 6 Spectrograms of 'better' (GA) and 'better' (RP)

The first 'better' is spoken in GA and the second is spoken in RP. The main difference between the two spectrograms is found in duration of the intervocalic /t/. In RP, the /t/ is clearly pronounced with the duration of 102 ms, but in GA, it is voiced and pronounced much shorter with the duration of 15 ms. Unlike aspiration, pronouncing the voiced /t/ as the normal /t/ does not give any effect semantically. Semantically they are irrelevant, and the normal /t/ is good enough. But if learners want to pronounce GA well, the voiced /t/ is one of the key features to make English sound like GA. In this respect, the voiced /t/ is also worth being emphasized. Some GA examples are shown as follows:

'pair', 'pare' [ˈpʰeɹ] 'tiger' [ˈtʰaɪgəɹ]

'vacation' [veɪˈkʰeɪʃn] 'city' [ˈsɪɾi]

'meeting' [ˈmiːɾɪŋ]

The fifth feature, which is newly added in this paper, is to display the pre-fortis clipping (e.g. [iˑ] and [ĭ]). Since the school textbooks adopt the quantitative approach, and quantity differences are important in Japanese, learners may fail to notice that vowel quantity is likely to change in accordance with the environment in English. This feature should be visually displayed. To illustrate how the pre-fortis clipping operates, spectrograms of 'tag' and 'tack' are displayed in Fig. 7.

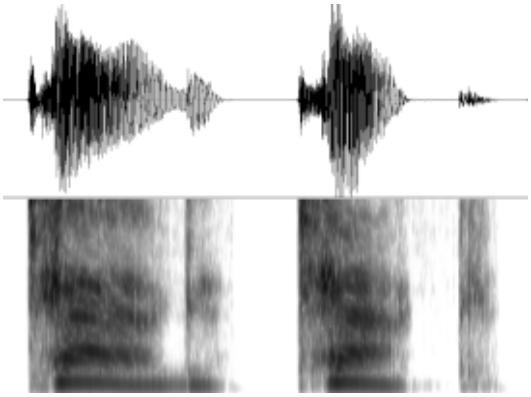

Fig. 7 Spectrograms of 'tag' and 'tack'

This figure shows that although there is almost no difference in word length (606 ms in 'tag' vs. 600 ms in 'tack'), such a difference is clearly existent in vowel length: 334 ms in 'tag' vs. 243 ms in 'tack'. In English, a word-final stop is not usually released, and in such cases, the key to distinguishing 'tag' from 'tack' is in vowel duration. Japanese learners of English have to be particularly sensitive to this difference, because no such difference is existent in Japanese. To illustrate this, 'kazu' ('number') and 'kasu' ('to impose') are acoustically analyzed in Fig. 8.

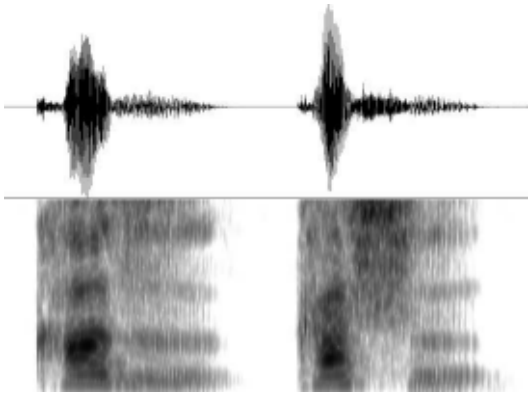

Fig. 8 Spectrograms of 'kazu' and 'kasu'

To avoid possible accentual influence, the two words have the same accent pattern — the head-high type. The first vowel /a/ is followed by a voiced consonant /z/ in 'kazu' and by a voiceless consonant /s/ in 'kasu', but the quantity difference in these two vowels is negligible: 105 ms in 'kazu' and 93 ms in 'kasu'. This clearly shows that the pre-fortis clipping does not work in Japanese. It should be noted here that this result has nothing to do with the quantity difference in the words because there is almost no difference in 'kazu' (406 ms) and 'kasu' (405 ms). It is true that the quantity difference plays an important part in semantic difference in Japanese, but this is realized as a double vowel (e.g. 'toshi' ('city') vs. 'tooshi' ('combative spirit) and 'toji' ('useless thing') vs. 'tooji' ('in those days')). In this way, the pre-fortis clipping is an entirely new feature to Japanese learners of English, and it is worth describing in IPA-J. Some GA examples are as follows:

'heart' [ˈhɑˑrt] cf. 'hard' [ˈhɑːrd]

'pat' [ˈpʰæ̆t] cf. 'pad' [ˈpʰæd]

'life' [ˈlă˕f] cf. 'live' [ˈla˕v]

cf. 'coffee' [ˈkʰɑˑfi] or [ˈkʰɑːfi]

cf. 'latter' [ˈlæ̰tɚ], 'ladder' [ˈlædɚ]

The pre-fortis clipping symbol is useful, especially when there is a contrast in minimal pairs, as in 'mat' vs. 'mad'. When there is no such contrast, as in 'coffee', however, this symbol may not be necessary in order to make transcription easier to read.

The way of treating the voiced /t/ is a little confusing when the pre-fortis clipping is involved. Takebayashi and Saito (1998: 88) state that when /t/ is pronounced as the voiced /t̬/, the pronunciation of 'ladder' and 'latter', of 'pudding' and 'putting', and of 'Adam' and 'atom' tends to be identical, but that because of the pre-fortis clipping, the vowel before /d/ is pronounced longer than the vowel before the voiced /t̬/. If this statement is strictly observed, the pre-fortis clipping symbol should be applied above /æ/ in the transcription of 'latter', but no such application is made. This is because among his data, the author did not find a noticeable quantity difference between such vowels. Fig. 9 illustrates this.

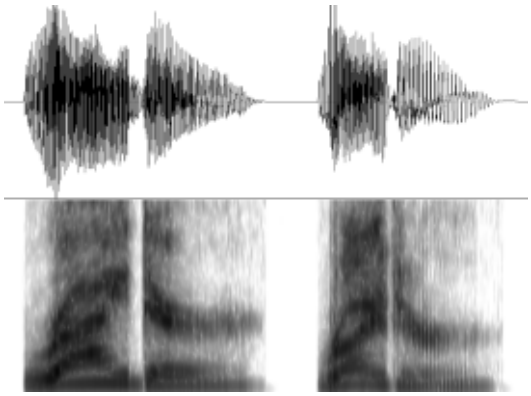

Fig. 9 Spectrograms of 'rider' and 'writer'

This figure shows that the diphthong in 'rider' is longer than the diphthong in 'writer': 243 ms vs. 179 ms. Some may feel that this difference supports Takebayashi and Saito's statement. The author, however, does not think that this is the right way of comparison, because word duration is also different. When a word is pronounced longer, segments inside the word, especially those in the stressed syllable are also pronounced longer. In this pair of examples, 'rider' is pronounced longer than 'writer': 712 ms vs. 524 ms. In cases like this, comparison should be made in terms of vowel-to-word duration. It is interesting to learn that vowel-to-word duration is identical between the two words: 34% (= 243/712*100) for 'rider' vs. 34% (= 179/524*100) for 'writer'. A similar result comes out when 'ladder' and 'latter' are compared in terms of vowel-to-word duration: 40% (= 189/477*100) for 'ladder' vs. 36% (= 195/535*100) for 'latter'. Unlike Takebayashi and Saito's statement, it seems that the effect of the pre-fortis clipping is weakened or lost when the /t/ in question becomes the voiced /t/. Getting back to Fig. 6, where 'better' is pronounced in GA and RP, it is known that the pre-fortis clipping is influential in RP. The vowel-to-word duration is different: 34% (= 150/443*100) in GA vs. 26% (= 93/354*100) in RP.

Some examples of IPA-J with all the five features are shown as follows. This time RP examples are also transcribed side by side for reference.

'motto' [ˈmɑːɾtoʊ | ˈmɒtəʊ (ˈmɒtəʊ)]

'entertain' [ˌɛnɾəˈtʰeɪn, ˌɛnə- | ˌɛntə-][13]

'community' [kəˈmjuːnəɾi | -nəti]

'university' [ˌjuːnɪˈvɜˑsəɾi (-ˈvɜːˑ-) | -ˈvɜˑsəti (-ˈvɜː-)]

'concentration' [ˌkʰɑːnsnˈtreɪʃn | ˌkʰɒn-]

'current' [ˈkɜːənt | ˈkʌrənt]

6. Approximate Kana Transcription

Since its debut, AKT has been improved many times, but it is still not problem-free. The author would like to make some comments, including two constructive criticisms here.

It is difficult to transcribe English phonemes in kana mainly for two reasons. First, it is basically impossible to transcribe consonant clusters and word-final consonants in kana. In English, it is possible to have three consonants in the onset and four consonants in the coda, and there are many closed syllables. These features are non-existent in Japanese. Except for '/j/ after specific consonants like /k/, there are no consonant clusters in Japanese. In addition to this, except for 'ン' (i.e. /m, n, ŋ, N/, depending on the phonetic environment), kana always includes a vowel. Second, it is basically impossible to transcribe sounds of another language with those of a mother tongue when the former language has much a richer phonological system than the latter. The number of phonemes is much greater in GA than in Japanese: 17 vowel phonemes in GA vs. five in Japanese. This difference — more than three times larger — is too great to be ignored. It seems, however, that there is logical adequacy in AKT because it aims not at transcribing model English sounds but their approximation by showing direction or vector to reach them.

There is, however, an advantage in AKT when it effectively makes use of the Japanese open syllable. In English, a word-final consonant and a word-initial vowel tend to be linked. In 'read it again' and 'put it on', for example, all three words are linked as if they were one word. Except for the last word, there are no closed syllables in these phrases. AKT is very useful in transcribing this linking. This feature reminds the author of an episode that he heard from Shimaoka. During one of his classes, a student could not pronounce 'Take it easy' well, but when he wrote this pronunciation in kana on the blackboard, he was able to pronounce it pretty well. This episode prompted him to come up with the idea for AKT. In addition to the effective use of the kana transcription, Shimaoka's teaching experience may have helped this student to pronounce this phrase well. There is one more advantageous condition involved in this particular phrase. The final word ends in a vowel. This is an ideal example in which kana works perfectly.

There are some kana scripts that worry the author. Two examples are mentioned here. The first is 'ヷ'. This is created by adding a small circle[14] to the right shoulder of 'ウ' (i.e. /u/). The reasoning behind the invention of this nonexistent kana may be as follows. In Japanese, /v/ is

sometimes transcribed as 'ヴ' by adding a double grave accent mark[15] to the right shoulder of 'ウ', as in 'ヴァイオリン' ('violin'). If the double grave accent mark means a voiced consonant and the small circle means a voiceless consonant (e.g. 'バ' (i.e. /ba/) vs. 'パ' (i.e. /pa/)), it is assumed that /f/ is transcribed by adding the small circle to 'ウ'. However, there are four problems with this assumption. First, very few Japanese would pronounce 'ヴ' as /v/. This consonant is not existent in the Japanese language. To have learners pronounce 'ヴ' as /f/, it must be presupposed that they pronounce 'ヴ' as /v/. Second, the meaning of the small circle is inconsistent. It is only in the /pa/ and /pja/ columns that this symbol is used to mean a voiceless consonant. In the /ka/, /sa/, /ta/ and /ha/ columns, no symbol is added to mean this. The /ha/, /ba/ and /pa/ columns comprise a complicated system. Consonants in the /ha/ column are fricatives, while those in the /ba/ and /pa/ columns are stops. If kana scripts in the /ha/ column were entirely different from those in the /ba/ and /pa/ columns, it might be helpful to transcribe /v/ and /f/. Third, 'ウ' is a vowel, not a consonant. If 'ヴ' and 'ヷ' are used as consonants, it is more educational to use /v/ and /f/. Fourth, it is difficult to find a phonetic similarity between 'ヷ' and /f/, not to mention 'ヴ' and /v/.[16] This phonetic similarity is important to enable learners to associate the target sound with its approximation.

The second kana script that worries the author is the use of the superscript 'ヌ' to signal the contact of the tongue tip against the alveolar ridge in pronouncing /l/. 'Like', for example, is transcribed as 'ˢライㇰ' (Shimaoka 2005). Since this kana script is composed of a consonant /n/ and a vowel /u/, the expected contact will be lost during the pronunciation of the vowel. It appears that 'ヌ' is used to mean only /n/ in AKT, but if that is the case, /n/ is more educational. It is also true, however, that this /n/ is redundant information when /l/ is used in the transcription.

Traditionally, the kana transcription has been used in many English-Japanese dictionaries, especially for beginners. Unfortunately, there is no unified kana transcription in Japan. The author thinks that AKT is one of the best kana transcription systems ever invented, but because of the great differences in the phonological systems and the writing systems of English and Japanese, he wonders whether kana is really useful in teaching English pronunciation to Japanese learners of English.[17]

7. Korean IPA Transcription

Lee attempts to use the Korean letters to replace the phonemic symbols in transcribing English. This is not directly related to Japanese learners of English, but a brief mention is made of this system because the idea behind it is the same as the one behind AKT.

In transcribing English word-final consonants, Korean is more advantageous than Japanese because this language has closed syllables and its writing system is based upon phonemic transcription, just like the Japanese writing system in the Roman alphabet. Korean, however, has two major weak points in transcribing English. First, some English sounds are not existent in Korean, such as /f/ and /z/. As Lee (2004: 113-119) states, not all of the English phonemes can be transcribed with the Korean letters. Second, there is no voicing contrast in Korean. Instead, this language has a three-way distinction: weakly aspirated consonants, strongly aspirated voiceless consonants, and glottalized consonants (Clark and Yallop (1995: 101)). The weakly aspirated consonants can be voiced or voiceless depending on the environment. They become voiceless in a word-initial position, but voiced in an intervocalic position.

The 24 English consonants are described in the Korean letters as follows:

/p/ = <ㅂ>, /t/ = <ㄷ>, /k/ = <ㄱ>, /b/ = <?>, /d/ = <?>, /g/ = <?>,

/tʃ/ = <ㅈ>, /dʒ/ = <?>, /f/ = <?>, /θ/ = <?>, /s/ = <ㅅ>, /ʃ/ = <?>,

/h/ = <ㅎ>, /v/ = <?>, /ð/ = <?>, /z/ = <?>, /ʒ/ = <?>, /m/ = <ㅁ>,

/n/ = <ㄴ>, /ŋ/ = <ㅇ>, /r/ = <ㄹ>, /l/ = <?>, /j/ = <?>, /w/ = <?>

The question mark indicates consonants that cannot be described in the Korean letters. To overcome this situation, Lee (2004: 118) makes some rules to modify the letters, and there are four of them relevant to English. The first is to add the voice bar to a letter in order to derive voiced symbols from voiceless ones. <ㅸ> (i.e. /b/), for example, is derived from <ㅂ> (i.e. /p/). The second is to add a small circle under or over a plosive or affricate symbol to make its homorganic fricative symbol. <ㆆ> (i.e. /θ/), for example, is derived from <ㄷ> (i.e. /t/). The third is to delete a stroke from a plosive symbol to make its homorganic fricative symbol. <ㅍ> (i.e. /f/), for example, is derived from <ㅂ> (i.e. /p/). The fourth is to add a stroke to a trill symbol to make its homorganic lateral symbol. <ㄺ> (i.e. /l/), for example, is derived from <ㄹ> (i.e. /r/). Applying these rules to the above transcription, the Korean IPA transcription system is completed as follows:

/p/ = <ㅂ>, /t/ = <ㄷ>, /k/ = <ㄱ>, /b/ = <ㅸ>, /d/ = <ㄸ>, /g/ = <ㄲ>,

/tʃ/ = <ㅈ>, /dʒ/ = <ㅉ>, /f/ = <ㅍ>, /θ/ = <ㆆ>, /s/ = <ㅅ>, /ʃ/ = <ㅆ>,

/h/ = <ㅎ>, /v/ = <ㅱ>, /ð/ = <ㆅ>, /z/ = <ㅿ>, /ʒ/ = <ㆊ>, /m/ = <ㅁ>,

/n/ = <ㄴ>, /ŋ/ = <ㅇ>, /r/ = <ㄹ>, /l/ = <ㄺ>, /j/ = < ᆞ >, /w/ = <ㅜ>

With the help of these four rules, all 24 English consonants are transcribed with the Korean letters.

It seems that the Korean IPA transcription system is more logical and easier to understand than AKT, because one-to-one correspondence between the English consonants and the Korean consonantal letters[18] is possible. The inclusion of the IPA in the name of this Korean transcription system reflects Lee's attempt to replace completely the IPA with the standard and modified Korean letters. One important condition, however, is required to implement this system to Korean learners of English. They need to be familiar with the concept of voicing because there is no voicing contrast in Korean, unlike English and Japanese. In other words, they must be able to pronounce voiceless and voiced consonants and hear the difference accurately before they use this transcription system. If this condition is fulfilled, this transcription system will be useful to Korean learners of English as a first step to learning English pronunciation.

8. Concluding Remarks

This paper argues effective ways to improve the pronunciation of Japanese learners of English, particularly its segments, from the viewpoint of transcriptions. There are two main types of transcription: the IPA and kana. It goes without saying that the IPA is the best transcription. Among the three approaches in the IPA transcription, the quantitative-qualitative approach is optimum and most educational. It is fortunate that many English-Japanese dictionaries adopt this approach, not to mention Roach et al. (2003) and Wells (2000). In particular, the method of transcription adopted in Roach et al. (2003) is better and easier to read. Syllabification is also problematic, though it has partly been dealt with in this paper. Roach et al. (2003) and Wells (2000) adopt different methods of syllabification, but it is not easy to conclude which is better. Unlike the current trend in many English-Japanese dictionaries, English school textbooks approved by Japan's Ministry of Education still use the traditional quantitative approach. As explained in this paper, this approach may be highly misleading to Japanese learners of English and should be abolished and changed to the quantitative-qualitative approach as soon as possible.

For those Japanese learners of English who are not familiar with the IPA, the author presents the visually modified IPA, called 'IPA-J' (i.e. the IPA for the Japanese). This is a broad phonetic transcription and has been invented by the author in order to add to the IPA selected phonetic information that may be necessary for Japanese learners of English. This information is selected through the author's teaching experience and his comparative study of English and Japanese. In this paper, five features of IPA-J are pointed out. The first feature is to make the first element of a

diphthong more noticeable by displaying it larger than the second. The second feature is to make a strong syllable more noticeable by displaying it larger, in bold type and making it shaded or colored. The third is to place stress marks before stressed syllables. The fourth is to display aspiration and the voiced /t/. The fifth feature is to display the pre-fortis clipping. There may be other features that Japanese learners of English find necessary or there may be better ways to display these five features. It may also be necessary to conduct experiments to learn whether IPA-J is really effective in teaching Japanese learners of English. The author would like to tackle these problems one by one to make IPA-J better and more effective. IPA-J should be used only for those who are unfamiliar to the IPA. Once the learners are accustomed to reading the IPA, its role is over. In this sense, IPA-J does not have to be applied to all words in dictionaries. Some basic words will do.

Another alternative transcription for Japanese learners of English is kana. There is no unified kana transcription in Japan. Among the various types, Approximate Kana Transcription, called AKT, is one of the best methods. There are many unique ideas involved in this transcription, and the author believes that there are many people who support AKT. Because of differences in the phonological systems and the writing systems of English and Japanese, however, there are difficult problems to solve. In addition to AKT, the Korean IPA transcription is also presented in this paper. It is impossible to transcribe all the English consonants in this system, but with the help of Lee's rules, the one-to-one correspondence between the English consonants and the Korean consonantal letters becomes possible. Basically, consonant letters are used in this system, which is more ideal than the approach adopted in AKT. There is, however, an important condition before this system is used properly. Unlike English and Japanese, Korean does not have voicing contrast. Korean learners of English must overcome this difficulty.

In order to master English pronunciation; sound, transcription and evaluation must be integrated effectively. The author believes that Japanese learners of English, irrespective of age, can acquire English pronunciation well enough, without traces of Japanese accent, if they are highly motivated, use effective transcription and get appropriate feedback from their teachers or technologically advanced PC programs.

In this paper, transcription of English segments is discussed, but prosody is also an important aspect. Some say that prosody is more important than segments to make oneself understood in English. The author believes that the tonetic stress mark system is the best approach, but in Japan, the so-called Pike system is still prevalent. This problem will be discussed in another paper.

(Professor, The Faculty of Economics, Takasaki City University of Economics)

References:

Clark, J. and Yallop, C. (1995). *An Introduction to Phonetics and Phonology.* Second Edition. Oxford: Blackwell.

Cruttenden, A. (2001). *Gimson's Pronunciation of English.* Sixth Edition. London: Arnold.

International Phonetic Association. (1999). *Handbook of the International Phonetic Association: A Guide to the Use of the International Phonetic Alphabet.* Cambridge: Cambridge University Press.

Ladefoged, P. (2001). *A Course in Phonetics.* Fourth Edition. Boston: Heinle & Heinle.

Ministry of Education. (1999a). *Course of Study for Junior High School Foreign Language Teaching (Chugakko Gakushu Shido Yoryo Kaisetsu Gaikokugohen).* Tokyo: Tokyo Shoseki.

Ministry of Education. (1999b). *Course of Study for High School Foreign Language (English) Teaching (Kotogakko Gakushu Shido Yoryo Kaisetsu Gaikokugohen Eigohen).* Tokyo: Kairyudo.

Lee, H. B. (2004). 'In Search of a Universal Phonetic Alphabet: Theory and Application of an Organic Visible Speech.' *Interspeech* 2004 ICSLP. 8th International Conference on Spoken Language Processing. pp. 113–120.

Roach, P., Hartman, J. and Setter, J. (eds.) (2003). *English Pronouncing Dictionary.* 16th Edition. Cambridge: Cambridge University Press.

Shimaoka, T. (2005). *Application of Approximate Kana Transcription: English-English-Japanese 2000 Word Dictionary.* Tokai: Ibaraki Insatsu.

Takebayashi, S. (1996). *English Phonetics (Eigo Onseigaku).* Tokyo: Kenkyusha.

Takebayashi, S. and Saito, H. (1988). *An Introduction to English Phonetics (Eigo Onseigaku Nyumon).* Second Edition. Tokyo: Taishukan Shoten.

Upton, C., Kretzschmar, W. A. and Konopka, R. (2001). *The Oxford Dictionary of Pronunciation for Current English.* Oxford: Oxford University Press.

Wells, J. C. (1990). 'Syllabification and Allophony'. In S. Ramsaran (ed.), *Studies in the Pronunciation of English: A Commemorative Volume in Honour of A. C. Gimson.* London: Routledge, pp. 76–86.

Wells, J. C. (2000). *Longman Pronunciation Dictionary.* New Edition. Essex: Pearson Education Limited.

Yomiuri Shimbun. (2006). 'English To Be Compulsory From Fifth Graders'. Tokyo: Yomiuri Shimbun, p. 1.

Yuzawa, N. (1999). 'An Exhortation to IPA-J' ('IPA-J No Susume'). *Eigo Kyoiku.* October Issue. Tokyo: Taishukan Shoten, pp. 42–43.

Yuzawa, N. (2000). 'IPA-J and DST' ('IPA-J To DST'). *English Pronunciation and Transcription (Eigo No Hatsuon To Hyoki).* Volume 2. English Phonetics and Transcription Association, pp. 60–81.

Notes:

1 Along this line, in 2000, the author gave a presentation on the importance of studying how much of a Japanese accent would be acceptable when speaking in English, but no further development has been made since then because of lack of funds and large data.

2 According to the Yomiuri Shimbun (March 28, 2006), English teaching will be implemented from the fifth grade in 2010.

3 The vowel insertion is a common error among Japanese learners of English.

4 Roach et al. (2003) and Wells (2000) adopt the quantitative-qualitative approach, but there is no consistency in Upton et al. (2001), where the quantitative approach is adopted for RP while the qualitative approach is adopted for GA.

5 This is known as the 'pre-fortis clipping' and is explained in Section 5.

6 When the author was asked by a school textbook publishing company to write pronunciation exercises, he proposed adopting the quantitative-qualitative approach, but this approach was rejected because it is not used in other textbooks and because it is not well known among teachers.

7 The problems of (1) and (7) were already discussed.

8 Shimaoka later introduced a new method of transcription called IPA+.

9 If color is available, it can be used instead of shade.

10 It should be noted here that the second element of the diphthong is not shown in a large font because

only the first element is made large. The second element is transcribed in a bold type and with shade.

11 In IPA-J, the syllabic consonant symbol is not used to make transcription easier to read.

12 The inferior vertical stroke indicates secondary stress.

13 The pre-fortis clipping symbol is not used in secondary stress to make transcription easier to read.

14 This small circle is called 'handakuten' in Japanese. It is used to show /pa, pi, pu, pe, po, pja, pju, pjo/ in Japanese. Sometimes, it is also used to show/ŋa, ŋi, ŋu, ŋe, ŋo/., but the use for these nasals is not standard.

15 This symbol is called 'dakuten' and means voicing.

16 As is discussed in the next section, /u/ is derived from /w/ in the Korean IPA transcription. These two segments are phonetically similar.

17 A former Japanese professor of English told the author that the IPA was thoroughly taught in the first two to three months in junior high school in the prewar Japanese education system. The author thinks that this style of education should be revived.

18 Exceptions are /j/ = < ᆡ > and /w/ = <ᅮ>. They are made by adding a hook to the vowel letters < ㅣ > and <ㅜ>, which mean /i/ and /u/ respectively. Since these two approximants and the corresponding vowels are phonetically similar, there is no logical problem in using these vowel letters to transcribe these two consonants.