

# 東京音楽大学リポジトリ

## Tokyo College of Music Repository

A Contrastive Study of the English and the  
Japanese Vowels and Consonants : Sounds with  
which the Japanese should take special care :  
日英母音及び子音の対照研究

メタデータ	言語: ja 出版者: 公開日: 1986-01-01 キーワード (Ja): キーワード (En): 作成者: Tamura, Kengo メールアドレス: 所属:
URL	<a href="https://tokyo-ondai.repo.nii.ac.jp/records/686">https://tokyo-ondai.repo.nii.ac.jp/records/686</a>

This work is licensed under a Creative Commons  
Attribution-NonCommercial-ShareAlike 3.0  
International License.



# A Contrastive Study of the English and the Japanese Vowels and Consonants

—Sounds with which the Japanese should take special care—

日英母音及び子音の対照研究

Kengo TAMURA

## Contents

### I Introduction

### II Vowels

- 1) Phonetic Symbols
- 2) English Pure Vowels
- 3) English Compound Vowels
- 4) Japanese Vowels
- 5) Comparison of Japanese and English Vowels
- 6) Vowels with which the Japanese should take special care  
(1) [ə] (2) [æ]

### III Consonants

- 1) English Consonants
- 2) Japanese Consonants
- 3) English Consonants with which the Japanese should take special care  
(1) [f, v] (2) [θ, ð]  
(3) [l] (4) [r]
- 4) English Consonant Clusters with which the Japanese should take special care  
(1) [tr-, dr-] (2) [st-, sw-, tj-]  
(3) [-pt, -kt] (4) [-tn, -tl]

### IV Conclusion

Bibliography

## I Introduction

In this paper, vowel sounds such as [ə] and [æ], consonant sounds such as [f, v], [θ, ð], [l], [r], and some consonant clusters which are difficult for Japanese learners to articulate, will be taken up and discussed from the viewpoint of pronunciation education. Japanese learners have great difficulty in correctly articulating these sounds mentioned above, because standard Japanese does not have these sounds.

In order to master English pronunciation, it would be ineffective and time-consuming for Japanese adults, who have already acquired the sound system of their mother language, merely to imitate and repeat the pronunciation of native speakers of English, although the imitation and repetition of the sounds will be necessary for infants who have not yet acquired the sound system of their mother language. Japanese learners would achieve good results immediately if they studied English pronunciation by comparing it with Japanese pronunciation. In this paper, English sounds are contrasted with Japanese sounds and the differences between them are discussed for the benefit of better English pronunciation.

It is to be hoped that Japanese learners will try to master English sounds, taking note of the differences between the two languages and making the best of the result of their study.

## II Vowels

### 1) Phonetic Symbols

There is no standard set of phonetic symbols; different scholars have adopted different conventions. In *An Outline of English Phonetics*, Daniel Jones recognizes [i:, i, e, æ, α:, ɔ:, ɒ, u:, ʊ, ʌ, ə:, ə] as English vowels and in *An Introduction to the Pronunciation of English*, A. C. Gimson recognizes [ɪ, e, æ, ɒ, ʊ, ʌ, ə] as short vowels and [i:, u:, α:, ɔ:, ɜ:] as long vowels. Robert L. Lado and Charles C. Fries recognize [i, ɪ, e, ε, æ, a, ɔ, ʊ, u, ə] in *English Pronunciation*, and John S. Kenyon and Thomas A. Knott recognize [i, ɪ, e, ε, æ, a, α, ɒ, ɔ, ʊ, u, ɜ, ɜ, ə, ə] in *A Pronouncing Dictionary of American English*. As shown in the previous paragraph, some scholars use the length mark to denote long vowels [i:, α:, ɔ:, u:, ə:], but some do not. In this paper, the phonetic symbols recognized by A. C. Gimson will be used, even though the phonetic symbols acknowledged by Daniel Jones have come into wide use in general dictionaries and

textbooks published in Japan. The reasons for selecting Gimson's phonetic symbols are as follows. Daniel Jones often used similar symbols to denote slightly different sounds which should be distinguished clearly from a phonetic viewpoint. For example, the difference between [i:] and [i] is not only the difference in length but also the difference in sound quality, so Gimson's [ɪ] is more suitable than Jones' [i]. In addition, the difference of sound quality between [u:] and [u] is remarkable, so [ʊ] is used here instead of [u]. Also the difference between [ə:] and [ə], and between [ɔ:] and [ɔ] is not only the difference of length but also the difference of sound itself. Therefore, Gimson's phonetic symbols are more appropriate for Japanese learners from the viewpoint of English education.

## 2) English Pure Vowels

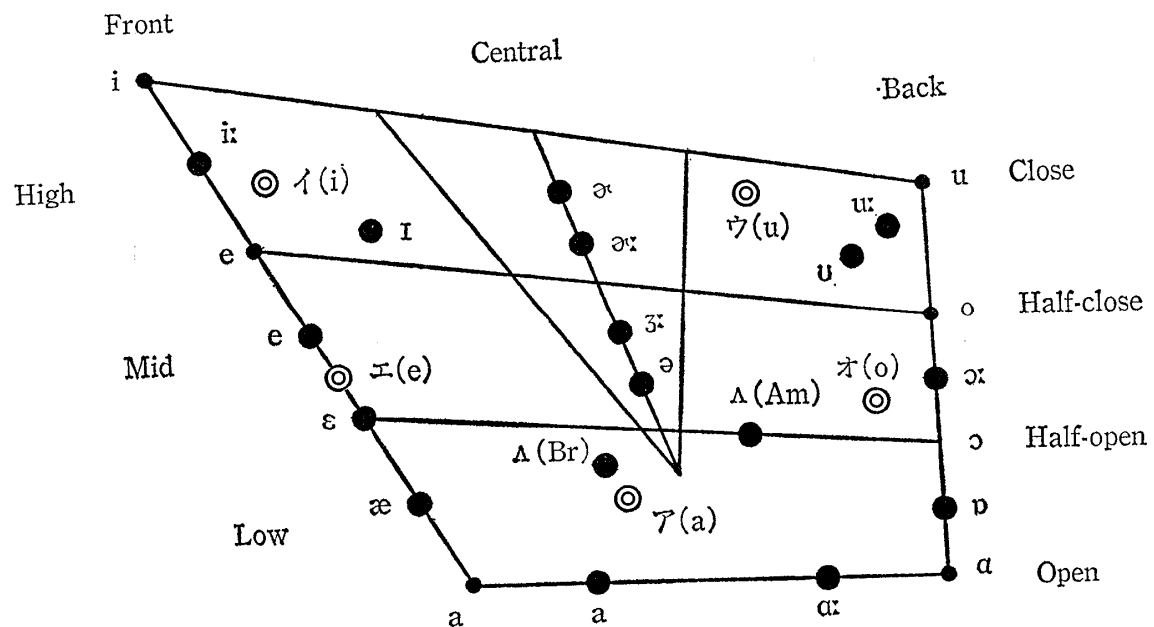


Fig. 1

The relationship of English Pure Vowels to the Cardinal Vowels

- Cardinal Vowel
- English Pure Vowel
- ◎ Japanese Vowel

The following points should be kept in mind when studying the preceding diagram :

(i) The back vowels are pronounced with rounded lips, and the degree of rounding increases markedly towards the high vowels.

- (ii) The muscles of the articulatory organs are more tense with [i:, u:, ɜ:] than with [ɪ, ʊ, ə].
- (iii) As for the length of pure vowels, low vowels are generally longer than high vowels. The length of vowels in front of voiced consonants is longer than that of vowels in front of voiceless consonants. The length of vowels without succeeding consonants is the longest.
- (iv) Vowels, when stressed, become long. Stressed vowels are longer than unstressed ones.
- (v) The vowels [ə:] and [ə] are used in General American speech.
- (vi) American phoneticians do not use the length mark.

**Table 1 The Length of Vowels**

Words	Length	Words	Length
si:	0.317 sec.	ka:	0.406 sec.
si:d	0.252 "	ka:d	0.338 "
si:n	0.199 "	ka:m	0.305 "
si:t	0.124 "	ka:t	0.254 "
si:tiŋ	0.087 "	ka:tiŋ	0.187 "

### 3) English Compound Vowels

A compound vowel is not a single sound but a combination of vowels, uttered in rapid succession, which forms a glide within one syllable. One of the component vowels is more strongly stressed than the others. The compound vowel is also known as a kinetic vowel.

There are two types of compound vowels: diphthongs and triphthongs. A diphthong is not a mere succession of two vowels. In order to be a diphthong, the two vowels must form a monosyllable, that is, the sonority of one of the two vowels must form the syllabic element by being stronger than the other, and the vowel with the weaker sonority must combine closely with the syllabic element, as a non-syllabic element.

A diphthong, a blend of two vowels, is a glide from one vowel to another. The syllabic element in a diphthong is the starting point of the glide. When pronouncing a diphthong, the non-syllabic element should be viewed as target, and not as the culminating point of the glide. If the non-syllabic element is viewed simply as a culminating point, the diphthong tends to be pronounced as two separate vowels. A diphthong should be

pronounced as though the non-syllabic element were the intended target, but this target is not reached.

Diphthongs are classified into falling diphthongs and rising diphthongs, according to the position of the syllabic element. All the English diphthongs are falling diphthongs except for a few unimportant exceptions.

There are nine English diphthongs. [eɪ], [aɪ], [ɔɪ], [əʊ] and [aʊ] which move in the direction of [ɪ] and [ʊ] are rising diphthongs, because the tongue is elevated toward the high position. [ɪə], [eə], [ʊə] and [ɔə] which move in the direction of [ə] are centring diphthongs.

Diphthongs are usually longer before voiced than before voiceless consonants, and those in the final position are the longest, e. g. *high* > *hide* > *height*. Most of the length and stress are concentrated on the first element, the second element being only lightly sounded, e. g. *day* ['deɪ], *made* ['meɪd]. A stressed diphthong is longer than an unstressed diphthong. Diphthongs also tend to be longer when they appear at the end of a phrase or sentence before a pause; [ɔɪ] of “*Look at that boy* [bɔɪ].” is longer than [ɔɪ] of “*Look at that boy’s* [bɔɪz] *hat.*”

The most common English triphthongs are [aɪə] (*fire, lion*) and [aʊə] (*tower, power*). In each of these triphthongs, it is natural that the first element should be pronounced strongly and clearly, the second element shortly and very weakly, and the last element a little longer than the first and second elements. In British pronunciation, triphthongs are occasionally pronounced indistinctly, using long vowels, e. g. *fire* [fa:ə], *flower* [fla:ə]. The Japanese language does not have compound vowels similar to English ones. The Japanese pronounce every syllable very clearly and tend to articulate each component vowel distinctly, with the same length and strength.

#### 4) Japanese Vowels

The Japanese language has only five vowels: [i, e, a, o, u]. In *Japanese Phonology*, Han says that under the same conditions, the length of the vowels decreases in the following sequence, with [u] the longest: [u, i, o, e, a]. Also Japanese syllables have almost the same length and are pronounced distinctly. The double vowels such as [aa, ii] are pronounced [a:, i:] when there is no plus hiatus between them.

At this stage, it should be pointed out that after [w], only [a] is possible, and that after [y], only [a, u, o] is possible.

## 5) Comparison of Japanese and English Vowels

### (i) English [ɪ, i:] and Japanese [i]

As Fig. 1 on page 6 shows, when we pronounce the English [ɪ], the highest point of the tongue is a little lower and a little further back than in the English [i:] and Japanese [i]. [ɪ], which is a lax vowel, is a sound between the Japanese vowels [i] and [e], and when not stressed, it is pronounced as an obscure vowel [ə].

In American English, [ɪ] doesn't appear at the end of the word, so "*Mary and city*" will be pronounced something like [ˈmæəri:] or [ˈsɪti:].

We can substitute the English [i:] for the Japanese [i:], but we cannot substitute [ɪ] for the Japanese [i].

### (ii) English [ʊ] and Japanese [u]

When pronouncing the English [ʊ], the tongue and the lips should be relaxed and the lips should be protruded. The position of the tongue is higher than that when pronouncing [ɔ]. The Japanese [u] is articulated without rounded lips unlike the English vowels [ʊ, u:], and the upper and lower teeth are very close together, and the point of articulation of the tongue is further towards the front, making [u] almost a central vowel. The English [ʊ] is a lax vowel and the point of articulation is a little lower and a little further back than in the Japanese [u].

### (iii) English [ɒ] and Japanese [o]

The highest point of the tongue in the English [ɒ] is much lower and a little further back than in the Japanese [o].

The English [ɒ] is pronounced from the back of the mouth, opening the mouth wider than in the Japanese [o], rounding the lips and lowering the tongue.

### (iv) English [ʌ] and Japanese [a]

The British English [ʌ] is pronounced by opening the mouth a little wider than in the Japanese [a], but the positions of the tongue in the British [ʌ] and the Japanese [a] are almost the same. As Fig. 1 shows, the position of the tongue in the American pronunciation of [ʌ] is a little higher and a little further back than in the British [ʌ].

## 6) Vowels with which the Japanese should take special care

### (1) [ə]

Japanese learners should be very careful of this sound, because the Japanese language has no sound similar to this.

The central part of the tongue is raised to a point midway between half-close and half-open. The lips are a little spread and the muscles of the tongue and lips are relaxed. [ɜ:] and [i:] are articulated with tensed lips, but for the vowel [ə], the lips are neither rounded nor tense.

The vowel [ə] occurs in initial, medial and final positions of words, e. g. *about*, *company*, *sofa*. [ə] occurs very frequently in unstressed syllables, so it is often called a weak vowel. It can be represented by various spellings, e. g. *china*, *possible*, *support*, *policeman*, *occasion*, *father*, *doctor*, etc. This vowel is particularly short, weak and indistinct when it is not final, e. g. in *above*, *away*, *fatigue*, *afternoon*, but in final positions, e. g. in *better*, *collar*, *mother*, *over*, *China*, this vowel sounds more like [ɜ:] or [ʌ], although it is not usually so clear. All English vowels, when unstressed, are articulated as [ə] or a similar sound, however [i:] becomes [ɪ], and [ɪ] does not change except for a few cases. In general, those whose native language is not English tend to use strong vowels instead of using [ə], and consequently stress some words and syllables improperly. This vowel is one of the most frequent sounds in English, so it is essential for us to pronounce [ə] or [ə:] properly. Especially, the Japanese tend to pronounce English vowels as they are written, instead of using [ə]. Also, they are likely to use [ɑ:] instead of [ə:].

In *An Introduction to the Pronunciation of English*, A. C. Gimson wrote :

“In particular, the learner should note those syllables of a word containing /ə/, remembering that /ə/ is a sound which occurs very frequently in English and that correct obscuration of the unaccented syllables of a word is as much a part of the word’s accentual pattern as the stress expended on the accented syllables. In this connection, the learner may gain greater familiarity with the occurrence of /ə/ by reading English texts transcribed phonetically and by himself making a phonetic transcription of connected English.”

### (2) [æ]

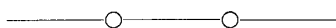
Standard Japanese, French and German do not have this sound, which is peculiar to English, so the Japanese are not good at articulating this sound correctly. They tend



to confuse this sound with [a] or [ɛ].

As Fig. 1 shows, in forming [æ], the tongue is low and at the front of the mouth, and occupies a position which is roughly intermediate between the positions for [ɛ] and [a]. The front and sides of the tongue are raised to just below the half-open position, with the side rims contacting the back upper molars slightly. The lips are spread or neutral, the opening between the jaws is medium to wide, and the mouth is rather more widely opened than for [e].

This sound comes between the Japanese [a] and [e]; a good way to approximate it is to try to pronounce the Japanese [a], while shaping the mouth for the Japanese [e]. [æ] is often produced with considerable constriction in the pharynx. This pharyngeal contraction is too vague to define precisely, so it may be useful for Japanese learners to tighten the throat in order to improve their pronunciation of this sound. Singers commonly use a modified [æ], or substitute [a] for it, because this sound cannot be sung beautifully. This vowel is a sort of strangled sound, and sounds as if it was being squeezed out. In order to obtain this sound, try to imitate the bleating of a sheep. Vowels, including [æ], are usually longer before lenis consonants [b, d, m] (*cab, mad, jam*) than before fortis consonants [p, t, k] (*cap, bat, back*). This vowel [æ], when stressed or put before the lenis consonants, is equivalent in length to the long vowels [i:, ɑ:, ɔ:, u:, ɜ:]. This vowel occurs in the initial and medial positions of words but never occurs in the final position, e. g. *apple, add, cat, black*. [æ] is almost always spelled with the letter *a* as in *bad, glad, lamp*. Other spellings for this sound are very exceptional: *au* in *aunt* and *laugh* in American and Northern English pronunciation and *ai* in *plaid, plait*.



(1) Daniel Jones: *The Phoneme*. 1950. § 403, 404

### III Consonants

#### 1) English Consonants

Vowels are voiced sounds and do not require any closure or narrowing in the speech tract. They do not cause any audible friction. On the other hand, consonants are characterized by a complete or partial closure of the breath stream, through the mouth

or nose. The English language has fourteen voiced consonants, and nine voiceless consonants.

Voiceless consonants tend to be lenis or weak sounds. Each consonant is classified according to its place and manner of articulation, and to the presence or absence of voice.

**Table 2. Classification of English Consonants**

Place of Articulation Manner of Articulation	Bilabial		Labio-dental		Dental		Alveolar		Alveo-palatal		Palatal		Velar		Glottal	
	vl	vd	vl	vd	vl	vd	vl	vd	vl	vd	vl	vd	vl	vd	vl	vd
Plosive	p	b					t	d					k	g		
Affricate									tʃ	dʒ						
Fricative			f	v	θ	ð	s	z	ʃ	ʒ						h
Nasal		m						n						ŋ		
Lateral								l								
Frictionless Continuant or Glide										r						
Semivowel										(r)		j		(w)		

vl=voiceless    vd=voiced

There are six plosive consonant phonemes in the English language as shown in Table 2. Plosive consonants are formed by completely stopping the breath at some point in the mouth, then holding or compressing the breath and abruptly forcing out the compressed air. When the fortis series [p, t, k] are initial in a stressed syllable, they are usually accompanied by a noticeable puff of air or aspiration, e. g. *pen* ['pʰen], *ten* ['tʰen], *kin* ['kʰɪn]. When [p, t, k] are preceded by [s] initially in the same syllable, there is no aspiration, even when followed by a stressed vowel, e. g. *spider* ['spɪdə], *stop* ['stɒp], *skit* ['skɪt]. When [p, t, k] are followed by an unstressed vowel or when they are in the final position, they are unaspirated or weakly aspirated, e. g. *polite* [pə'laɪt], *tobacco* [tə'bækəʊ], *skip* [skɪp], *sit* [sɪt], *look* [lʊk].

Initially, before a stressed vowel, the [b, d, g] sounds, in such words as *big*, *dog* and *get*, are partially devoiced, unaspirated and relatively strongly exploded. Those in the final position are partially devoiced, e. g. *rub* [rʌb̥], *sad* [sæd̥], *dog* [dɒg̥]. When situated

between voiced sounds, [b, d, g] are also voiced, e.g. *rubber* ['rʌbə], *leader* ['li:də], *eager* ['i:gə].

In American pronunciation, when [t] is situated between vowels or is followed by unstressed vowels, it is pronounced [t̚] or [d̚], e.g. *butter*, *letter*.

There are two affricate consonant phonemes in the English language, as shown in Table 2. An affricate is a combination or compound of a plosive and a fricative. [dʒ] is partially devoiced in the final position.

There are nine fricative consonant phonemes in the English language as shown in Table 2. Fricative consonants are formed by using two closely posed articulators to narrow the breath at some point in the mouth. When situated between voiced sounds, [v, ð, z, ʒ] are voiced, e.g. *lover* ['lʌvə], *breathing* ['bri:ðɪŋ], *pleasure* ['pleɪzə], *razor* ['reɪzə], but they are normally partially devoiced when they appear initially or finally in a sentence or phrase, and when they precede voiceless consonants.

There are three nasal consonant phonemes in the English language as shown in Table 2. Nasal consonants are formed by blocking the mouth passage completely at some point in the mouth, the soft palate being in its lowered position so that the air is pushed out of the nose. Nasal sounds are voiced, but when they follow voiceless consonants in the same syllable, they are partially devoiced, e.g. *smoke*, *snake*. When the [m] sound immediately precedes a labio-dental sound [f, v], the front closure may be labio-dental [m̚] rather than bilabial, e.g. in *comfort* ['kʌmfət], *come first* ['kʌm̚ 'fɜ:st], *triumph* ['traɪəmf̚].

The articulation of [n] is liable to be influenced by that of the following consonants, e.g. when followed by [f, v], as in *infant* ['ɪnfənt] *invoice* ['ɪnvɔɪs], [n] is pronounced [n̚]; when followed by [θ, ð], as in *month* [mʌn̚θ], *on the* [ɔn̚ ðə], the dental [n̪] is used; when followed by [p, b, m], as in *ten people (boys, men)* [ten pi:pl—bɔɪz, men], [n] is pronounced [m̚]; when followed by [k, g], as in *ten cakes (girls)* [ten keɪks—gə:lz], [ŋ] is used instead of [n].

When the spelling “ng” (or “ngue”) is final in a word, it is always pronounced [ŋ], e.g. *sing* [sɪŋ], *bring* [brɪŋ], *tongue* [tʌŋ]. Inflectional endings do not change the [ŋ] form, e.g. *singing* ['sɪŋɪŋ], *bringer* ['brɪŋə], *tongues* [tʌŋz]. The spelling “ng” in the medial position as in *finger*, is pronounced [ŋg] unless it is in a compound word (*springtime* ['sprɪŋtʌɪm]) or in an inflected form (*hanger* ['hæŋə]). However, the comparative and superlative forms of *long*, *young* and *strong* are exceptions: *long* [lɔŋ], *longer* ['lɔŋgə], *longest* ['lɔŋgɪst], *young* [jʌŋ], *younger* ['jʌŋgə], *youngest* ['jʌŋgɪst], *strong* [strɔŋ], *stronger* ['strɔŋgə], *strongest* ['strɔŋgɪst].

In English, there are two semivowels [w] and [j]. They glide rapidly from positions of approximately [u:] and [i:] respectively to those of the following vowels, which are the syllabic sounds, e. g. *we* [wɪ], *yacht* [jɔt]. Semivowels are, in phonetic terms, vocalic, but they are treated within the consonant class. They can not be central in the syllable because of their gliding nature.

## 2) Japanese Consonants

Table 3. Classification of Japanese Consonants

Place of Articulation Manner of Articulation	Bilabial		Labio Dental		Dental		Alveolar		Alveo Palatal		Palatal		Velar		Glottal	
	vl	vd	vl	vd	vl	vd	vl	vd	vl	vd	vl	vd	vl	vd	vl	vd
Plosive	p	b			t	d							k	g		
Affricate					ts	(dz)			tʃ	(dʒ)						
Fricative	(Φ)				s	z			(ʃ)	(ʒ)	(ç)					h
Nasal		m						n						(ŋ)		
Flap								ɾ								
Semivowel		w								j						

( ) allophone

The aspiration of Japanese plosive sounds [p, t, k] is much weaker than that of English voiceless plosive sounds. The explosion of Japanese [b, d, g] are weaker than those of English ones, and the Japanese [b, d, g] are completely voiced. In English, some words end in plosive sounds, e. g. *cat* [kæt], *stop* [stɒp], *kick* [kɪk], but this never occurs in Japanese. The Japanese [t, d] appears only before [e, a, o], as in *te*, *ta*, *to*, *de*, *da*, *do*.

The affricate [ts] is used in [tsu], e. g. 月 [tsuki], and [dz] appears in [dza], [dzu], [dze], [dzo], e. g. 残念 [dzanneN], 全部 [dzembu], ずるい [dzurui], 雑布 [dzo:kiN]. When pronouncing the Japanese [tʃ], the point of articulation is a little different from that of English; in Japanese, the blade of the tongue touches the alveolar ridge, and the tip of the tongue does not touch the teeth except when articulating strongly. On the other hand, when pronouncing the English [tʃ, dʒ], the tip of the tongue touches

the back part of the alveolar ridge, and the lips are a little more rounded than in Japanese.

The fricative sound [ɸ] is one of the characteristic sounds of the Japanese language. The place of articulation is bilabial, e. g. 富士山 [ɸuʒisaN], 双子 [ɸutago], 深い [ɸukai].

The Japanese [s] is articulated by causing slight friction by gently pressing the blade of the tongue against the hard palate. The English [s] is articulated by causing stronger friction by pressing the tip and blade of the tongue against the alveolar ridge.

The Japanese sound [z], which is a voiced [s], is a weaker sound than the English [z], and it appears in medial or final positions of words, e. g. あざみ [azami], まずい [mazui], 風 [kaze], 数 [kazu]. This sound should not be confused with [dz]. (cf. p 11)

The Japanese sound [ʃ] is almost the same sound as the English [ʃ], but the English [ʃ] is articulated with slightly more rounded lips. [ʒ] is a voiced [ʃ].

[ç] is another characteristic Japanese sound, and is articulated by causing friction by pressing the front of the tongue against the hard palate. This allophone appears only in [çi]. The Japanese [çi] must not be confused with the English [hɪ]. The English [h] is a voiceless glottal fricative sound. However, in standard English, [çu:] as well as [hju:] is possible as in *huge* [hju:dʒ / çu:dʒ], *human* [hju:mən / çu:mən], *hue* [hju: / çu:].

In Japanese, [n] is pronounced [m], when followed by [p, b, m] as in 散歩 [sampo], 勤勉 [kimbeN], 本命 [hommei]. In this case, [n] is pronounced without letting the tongue touch any part of the mouth. [n] is pronounced [n] when followed by [t, d, z, tʃ, n, ɾ] as in 簡単 [kantaN], 寛大 [kandai], 困難 [konnaN], 真理 [ʃinfi]. [n] is pronounced [ŋ] without letting the tongue touch any part of the mouth when followed by [k, g, ŋ] as in 印鑑 [inʒkaN], 金言 [kingeN].

The English [w] can precede almost all the simple and compound vowels and it is pronounced with very strongly rounded lips, but the Japanese [w] can only precede [a] as in 若い [wakai], and the lips are only slightly rounded. The English [j], in initial positions, can precede all the short and long vowels, but the Japanese [j] can only precede [a, u, o], e. g. 約束 [jakusoku], 百合 [juʃi], 呼ぶ [jobu].

### 3) English Consonants with which the Japanese should take special care

#### (1) [f, v]

Japanese learners should pay particular attention to the sounds [f, v], because these sounds are not native to the Japanese language.

[f] is a voiceless labio-dental fricative as shown in Table 2. This sound is formed by

gently pressing the lower lip against the edge of the upper front teeth and causing slight friction between them; the soft palate is raised so that no air passes through the nose and there is no vocal cord vibration. The actual point of contact between the lower lip and the upper teeth will vary somewhat according to the adjacent sound, e. g. if the adjacent sound is a back vowel such as [u:], the contact on the lower lip tends to be more retracted than for a front vowel such as [i:].

The Japanese generally tend to replace [f] by the bilabial fricative [Φ], *four* [fɔə] → [Φɔə], *roof* [ru:f] → [ru:Φ].

It would be very effective for Japanese learners to practice saying [fa, fi, fu, fe, fo] repeatedly. [v] is formed like [f] except that there is vocal cord vibration, so it is defined as a voiced labio-dental fricative consonant.

In order to pronounce [v] properly, at first the lower lip and upper teeth are gently pressed together and breath is forced out between them to form [f]. Then the vocal cords are vibrated to form [v]. [f] is a strong and long consonant, whereas [v] is weak and short. This is shown in the words *life* [laɪf] and *live* [laɪv]. But in [laɪf] the [aɪ] is quite short and in [laɪv] it is comparatively long. The Japanese generally replace [v] by the voiced bilabial plosive [b], e. g. *violin* [vaɪə'li:n] → [baɪə'li:n], *vitamin* ['vɪtəmi:n] → ['bɪtəmi:n]. Natives of London or other districts also replace [v] by [b], e. g. *seven* [sevn] → [sebm], *eleven* [ɪ'lebn] → [ɪ'lebm], *heaven* [hevn] → [hebm].

It would be very effective for Japanese learners to repeat [va, vi, vu, ve, vo] for practice.

## (2) [θ, ð]

Japanese learners should pay particular attention to the sounds [θ, ð], because these sounds are not native to the Japanese language.

[θ] is a voiceless dental fricative, as shown in Table 2. [θ] is formed by lightly contacting the inner surface of the upper teeth with the tip of the tongue and causing slight friction. The soft palate is raised so that all the breath is forced to pass through the mouth and there is no vocal cord vibration. This sound is very difficult for Japanese learners to articulate properly, so at an early stage it would be advisable for them to start by protruding their tip of the tongue between the upper and lower teeth.

The lip position will depend upon the adjacent sound; if the adjacent sound is a front vowel such as [i:], the lips tend to be spread, but for a back vowel such as [u:] they are somewhat rounded. The Japanese tend to replace [θ] by the voiceless blade-alveolar fricative [s], e. g. *think* [θɪŋk] → [sɪŋk], *thank* [θæŋk] → [sæŋk], *something* ['sʌmθɪŋ] → ['sʌmsɪŋ], *both* [bouθ] → [bous]. The [s] sound is articulated by lightly contacting

the upper alveolar ridge with the tip and blade of the tongue, and making a strong friction. The teeth are close together, the air passage between the tip and blade of the tongue and the upper alveolar ridge is extremely narrow, the soft palate is raised, and there is no vocal cord vibration. The friction for [s] is much greater than for [θ], and the upper and lower teeth are close together.

The [θs] sound as in *fifths* [fɪfθs], *sixth* [sɪksθ] and *months* [mʌnθs] may sometimes be pronounced in rapid speech as [fɪfs] or [fɪfts], [sɪks] or [sɪkts], [mʌns] or [mʌnts].

Some non-native speakers of English replace [θ] by [t], e.g. *three* [θri:] → [tri:], *thing* [θɪŋ] → [tɪŋ], and English infants and Cockneys tend to replace [θ] by [f], e.g. *thick* [θɪk] → [fɪk], *thirty* ['θɜ:tri] → ['fɜ:tri], but this is not standard English. The Japanese should avoid this kind of pronunciation. It would be very effective for Japanese learners to practice saying [θa, θi, θu, θe, θo] repeatedly.

[ð] is formed like [θ] except that there is vocal cord vibration, so it is defined as a voiced dental fricative. [ð] is the voiced cognate of [θ]. [θ] is a strong and long consonant, whereas [ð] is weak and short.

The Japanese generally replace [ð] by either [z] or [dz], e.g. *that* [ðæt] → [dzæt], *they* [ðeɪ] → [dzeɪ], *other* ['ʌðə] → ['ʌzə], *together* [tə'geðə] → [tə'gezə]. [z] is the voiced counterpart of [s].

Some non-native speakers of English replace [ð] by [d], e.g. *the* [ðə] → [də], *there* [ðeə] → [deə], *them* [ðem] → [dem], and Cockneys and English infants tend to replace [θ] by [v], e.g. *this* [ðɪs] → [vɪs], *then* [ðen] → [ven], but this is not standard English, so the Japanese should avoid this kind of pronunciation. This sound [ð] is often encountered in normal speech, so Japanese learners should be familiar with this sound and pronounce it correctly at an early stage.

It would be very effective for Japanese learners to repeat [ða, ði, ðu, ðe, ðo] for practice.

### (3) [l]

Japanese learners should pay particular attention to this sound because it is not native to the Japanese language. [l] is a voiced alveolar lateral consonant as shown in Table 2. [l] is formed by touching the teeth-ridge with the tip of the tongue in such a way that there is complete closure in the middle of the mouth, and passing the air out through one or both sides of the tongue. The soft palate is raised and there is vocal cord vibration. [l] has two clearly distinguishable allophones, commonly known as clear 'l' and dark 'l'. Clear 'l' occurs only before vowels and before [j], as in *leave*, *let*, *late*,

and *value*. It has the quality or resonance of a front vowel approximating to [ɪ]. Dark 'ɹ' occurs before consonants, finally after vowels as in *help*, *milk*, *feel* and *cruel*, and as a syllabic sound following a consonant, as in *apple* and *middle*. It has the quality or resonance of a back vowel approximating to [ʊ]. This dark 'ɹ' is transcribed phonetically as [ɹ̥]. The [ɹ̥] is a clear 'ɹ', when it is in the final position and immediately followed by a vowel at the beginning of the next word, e. g. *feel it*, *all over*, *fall out*. For clear 'ɹ', the front of the tongue is raised in the direction of the hard palate, while for dark 'ɹ' the front of the tongue is somewhat depressed and the back of the tongue is raised in the direction of the soft palate.

When the tip of the tongue is forcibly released from the alveolar ridge to make clear 'ɹ', the [ɹ̥] becomes very clear, but when it is released too forcibly, the [ɹ̥] becomes like a plosive sound approximating to [d], so the tip of the tongue should be released very smoothly. When pronouncing dark 'ɹ', the breath is weaker than in clear 'ɹ'. The dark 'ɹ' is likely to blend into other consonants easily, and the sound becomes very unclear and dark. [ɹ̥] in the final position is articulated similarly to the [ŋ] sound; the tip of the tongue should not be released from the alveolar ridge. If the tip of the tongue is released, [ɹ̥] becomes the Japanese *flapped r* or the English [r]. The point of contact of the tongue for [ɹ̥] is influenced by the place of articulation of the following or preceding consonant. [ɹ̥] is dentalized when it precedes or follows [θ, ð] as in *health* [hɛɪ̯θ], *wealth* [weɪ̯θ], *will they* [wɪ̯ðeɪ], *a month late* [ə mʌnθ leɪ̯t], *with love* [wɪ̯ð lʌv]. [ɹ̥] is palatalized when it immediately precedes the [j] sound as in *million* ['mɪ̯ljən], *will you come* [wɪ̯jə kʌm]. [ɹ̥] is nasalized when followed by a nasal consonant as in *elm* [elm̩], *kiln* [kɪ̯ln̩]. As for the degree of darkness, [ɹ̥] following back vowels as in *hold* and *hole* tends to be darker than [ɹ̥] following front vowels as in *field* and *feel*, and the degree of darkness in the syllabic [ɹ̥] as in *people* ['pi:pl̩], *table* ['teɪb̩] and *eagle* ['i:gl̩], is very remarkable. The Japanese tend to pronounce [ɹ̥] by touching the tip of the tongue not against the normal position—the teethridge—but a little further back because of the influence of the Japanese *flapped r*, and they do not keep the tongue against the palate long enough.

It would be very effective for Japanese learners to repeat [la, li, lu, le, lo] for practice.

(4) [r]

Japanese learners should pay particular attention to this sound because it is not easy for the Japanese to distinguish this sound from [ɹ̥] or the Japanese *flapped r*. [r] is a voiced post-alveolar frictionless continuant. The American English pronunciation of this



sound can be treated as a semivowel. [r] is formed by curling the tip of the tongue back until it is pointing at, but not touching, the hard palate, and by passing voiced air between the tip of the tongue and the palate without friction. The soft palate is raised. The lips are rather rounded, especially when [r] is at the beginning of the words and the following vowel is a back vowel.

When pronouncing the Japanese *flapped r*, the tip of the tongue touches the hard palate or the alveolar ridge. The *flapped r* is transcribed as [ɾ] by the IPA. Not only Japanese but many speakers of Received English use a sound similar to the *flapped r*. It occurs chiefly in unstressed intervocalic positions, as in *very* ['veri] → ['veɾi], *sorry* ['sɔri] → ['sɔɾi], *period* ['piəriəd] → [piɾiəd].

*Flapped r* is made with only one single tap of the tip of the tongue against the hard palate or the alveolar ridge. The RP [r] occurs only before vowels, never before consonants, as in *red* [red], *write* [raɪt], and *roof* [ru:f]. In words such as *port* [pɔ:t], *learn* [lə:n], *farm* [fɑ:m], the letter *r* is not pronounced in RP. But the Irish, the Scots, and the Americans pronounce the letter *r*. Even in RP, when a word ending with the letter *r* is immediately followed by a word beginning with a vowel, the letter *r* is usually pronounced, as in *a pair of shoes* [ə 'peər əv 'fu:z], *never again* ['nevər 'ægen], *far away* ['fa:r 'əweɪ]. This is called the *linking r*.

Many native speakers of English usually add [r] to words ending in vowels when the following words begin with vowels, even when there is no letter *r* in the spelling. Thus *Africa and Asia* is often pronounced ['æfrɪkər ən 'eɪʃə] instead of ['æfrɪkə ən 'eɪʃə]. Other examples are *the idea of it* [ðɪ 'aɪdɪər əv ɪt], *I saw it* [aɪ'sɔ:r ɪt], *drama and music* ['drɑ:mər ən 'mju:zɪk], *the law of England* [ðə 'lɔ:r əv 'ɪŋɡlənd]. This is called *intrusive r*. A very large number of native speakers of English, educated as well as uneducated, use the *linking r* or *intrusive r*, but the latter is not standard English and Japanese learners should avoid this sort of pronunciation.

As mentioned before, the Japanese tend to confuse [r] with *flapped r* or [l]. Japanese learners should be careful not to let the tongue touch any part of the mouth when pronouncing the English [r] sound.

It would be very effective for Japanese learners to repeat [ra, ri, ru, re, ro] for practice.

#### 4) English Consonant Clusters with which the Japanese should take special care

Consonant clusters can generally be classified into initial, medial and final consonant clusters.

(1) [tr-, dr-]

There are no sounds similar to [tr] or [dr] in the Japanese language.

[tr] is formed by placing the tip of the tongue for [t] behind the alveolar ridge, on the front of the hard palate and by moving the tip of the tongue from the alveolar ridge, not too rapidly, into position to make the friction for [r]. Air is compressed by pressure from the lungs and there is no vocal cord vibration. The soft palate is raised and the centre of the tongue is hollowed. The [r] in [tr] is voiceless.

[dr] is formed like [tr] except that the vocal cords are vibrated. In [dr] too, the tip of the tongue is further back than usual for [d] and there is friction as the voiced air passes over the tip of the tongue for the [r]. [tr] and [tʃ] have similar sounds, and [dr] and [dʒ] also sound similar to each other. Japanese learners tend to put a vowel between [t] and [r], and between [d] and [r], as in *train* [torem], *try* [torar], *drive* [dorarv], *dream* [dori:m]. They should think of either [tr] or [dr] as not two sounds but as one sound.

(2) [st-, sw-, tj-,]

As for [st], the sonority of the first sound [s] is louder than that of the second sound [t], and [s] should be articulated as weak and short as possible. The vowel following [t] should be pronounced strongly. The Japanese should be careful not to put a vowel between [s] and [t], as in *spy* [supai], *stay* [suter]. They should start with [s] and halt it by raising the tip of the tongue to cut off the friction without putting a vowel after it. To pronounce [sw] correctly, there should be no pause between [s] and [w], and at first [s] is pronounced and at the same time the lips are rounded for [w], that is to say, [s] is labialized for [w].

The Japanese tend to replace [w] by [u], as in *sweet* [sui:t], *quiz* [kuiz], *twelve* [tuelv], instead of [swi:t], [kwiz] and [twelv]. Most Japanese and some British people too, tend to pronounce [tj] as [tʃ], as in *tulip* [ˈtʃu:lɪp], *Tuesday* [ˈtʃu:zdi], *tube* [tʃu:b] instead of [ˈtju:lɪp], [ˈtju:zdi], [tju:b]. They also pronounce *duty* [ˈdʒu:ti], *duke* [dʒu:k] instead of [ˈdju:ti], [dju:k].

In American English, [tju] and [dju] are pronounced [tu] and [du], as in *tune* [tu:n], *duty* [ˈdu:ti], *student* [ˈstu:dənt].

(3) [-pt, kt]

When one plosive consonant is immediately followed by another, the closure for the

second consonant is made while the closure for the first is still in position.

To pronounce [pt]; the lips are closed for [p] and air is compressed by pressure from the lungs. Then, with the lips still closed, the tip of the tongue is placed on the alveolar ridge for [t]. After this, the lips are opened to articulate [p], but there is no explosion of air because of the tongue closure for [t]. Almost at the same time, the tip of the tongue leaves the alveolar ridge to articulate [t] and the compressed air explodes out of the mouth. There is only one explosion for the two plosives, and the first plosive [p] is incomplete. This is called an incomplete plosion.

As for [kt], the back of the tongue makes the closure for [k], then the tip of the tongue is placed on the alveolar ridge for [t]. After this the back of the tongue is lowered to articulate [k], but there is no explosion of air because the closure of the alveolar ridge prevents the compressed air from bursting out of the mouth. Almost at the same time the tip of the tongue leaves the alveolar ridge to articulate [t] and the compressed air explodes out of the mouth. The first plosive [k] is an incomplete plosion. The ending of [kt] is, as it were, swallowed. Japanese learners should be careful not to put a vowel between [p] and [t], or between [k] and [t] as in *kept* [kept], *stopped* [stɒpt], and *liked* [laɪkt]. When a plosive is followed by the same plosive, there is only one explosion, but the closure is held for double the usual time, e.g. *black cat* [blæk:æt], *big ground* [bɪg:raʊnd], *hot tea* [hɒt:i:], *want to* [wɒnt:u:], *good day* [gʊd:ei].

#### (4) [-tn, -tl]

When a plosive is followed by nasal [n] or [m], the explosion of the plosive is not formed by the air escaping through the mouth but through the nose. The tip of the tongue is placed on the alveolar ridge for [t], then the soft palate is raised to shut off the nasal cavity. After this, the soft palate is lowered, so that the breath explodes out of the nose rather than out of the mouth, with the tip of the tongue firmly touching the alveolar ridge all the time. This explosion is known as nasal explosion.

Japanese learners tend to explode the breath out of the mouth rather than out of the nose, as in *eaten* [i:tən], *good morning* ['gʊdə'mɔ:nɪŋ], *topmost* ['tɒpməʊst] etc., instead of [i:tn], ['gʊd'mɔ:nɪŋ], ['tɒpməʊst].

As for [tl] the tip of the tongue is placed on the alveolar ridge [t] and the sides of the tongue firmly touch the sides of the palate. Then the sides of the tongue are lowered from the sides of the palate to articulate [l], keeping the tip of the tongue touching the alveolar ridge so that the breath passes out laterally. This is called lateral explosion.

[dl] is formed in the same way but voiced. The Japanese tend to explode [t] on the alveolar ridge instead of on the sides of the palate, as in *bottle* [bɒtəl], *middle* [mɪdəl]. It must be noted that when pronouncing [tl] or [dl], the tip of the tongue does not leave the alveolar ridge.

## IV Conclusion

In this paper, plosives, affricates, nasals, and semivowels have not been discussed in depth. This is not because these sounds are not important, but only because [f, v], [θ, ð], [l] and [r] were given higher priority.

As mentioned before, the Japanese tend to pronounce English consonants very weakly, in the same way as they pronounce Japanese consonants; this is because of the weakness of the Japanese consonant phonemes. Because of this tendency, native speakers of English can not always understand the words which the Japanese articulate. The most important thing in English pronunciation is to exaggerate the sounds. It goes without saying that plosives should be spat out as strongly as possible. English fricatives should be pronounced with stronger friction than Japanese fricatives and they should be articulated for a longer duration. It should be noted that Japanese learners must be taught to pronounce English consonants very strongly.

As for vowels, as mentioned before, the Japanese should be very careful of vowels such as [ə] or [æ] which standard Japanese does not have. Attention should also be paid to compound vowels. The Japanese tend to pronounce the component vowels of compound vowels distinctly, with the same length and strength. Diphthongs should not be pronounced as two separate vowels. The first vowel of a diphthong should be articulated very strongly and distinctly, and the second vowel should be very weak and short. In addition to the pronunciation of vowels and consonants discussed in this paper, English rhythm, stress and intonation are very important too from the viewpoint of pronunciation. Even if speakers pronounce each vowel or consonant properly, but the rhythm or intonation is incorrect, they cannot make themselves understood by native speakers of English.

The main difference of rhythm between English and Japanese is described below. All the Japanese syllables are open syllables except for a "sokuon" (double consonant) and a "hatsuon" (syllabic nasal). Each of them is pronounced with the same length and regularity, and vowels are articulated precisely and clearly. Japanese rhythm which is

based on the syllabic unit is called syllable-timed rhythm. In the English language, on the other hand, stressed syllables tend to occur at evenly spaced intervals in time, and the unstressed syllables between the stressed syllables are articulated quickly and indistinctly. It is very difficult for the Japanese to catch or to express these English unstressed syllables precisely. English rhythm, decided by the stress, is called stress-timed rhythm.

It is desirable for Japanese learners to get used to English stress-timed rhythm at an early stage and to avoid pronouncing English with the same syllable length and stress, in the same way as they pronounce Japanese.

David Abercrombie and Robert Lado demand that language teachers who are not native speakers of English should be able to pronounce English almost as well as native speakers. Teachers should make a strenuous effort to reach this goal.

However, Japanese teachers of English have great difficulty in mastering the art of English pronunciation, as well as other English language skills, especially as they had already fully acquired the sounds of their mother language before starting to learn English themselves. Even if they have no confidence in their ability to pronounce English, they should not be daunted. The most important point for them is to have a responsible attitude toward their students and enthusiasm as a language teacher. Japanese people are basically shy and passive as far as foreign language learning is concerned, so at an early stage, teachers should encourage learners to speak in a loud voice.

In addition, foreign teachers should study the sound system of the mother language of their students.

It is said that language starts with the ear, and Otto Jespersen said that language is primarily speech.

Both teachers and learners should keep it in mind that at the first stage of learning, correct pronunciation must be mastered before concentrating on the other language skills.

(本学教授=英語担当)

### Bibliography

- Abercrombie, David. 1967. *Elements of General Phonetics*. Edinburgh University Press
- Bronstein, Arthur J. 1960. *The Pronunciation of American English*. Prentice-Hall
- Fries, Charles C. and Agnes C. Fries. 1961. *Foundations for English Teaching*. Tokyo: ELEC
- Gimson, A.C. 1973. *An Introduction to the Pronunciation of English*. Edward Arnold
- Jones, Daniel. 1972. *Everyman's English Pronouncing Dictionary*. Everyman's Reference Library
- . 1960. *An Outline of English Phonetics*. London. Cambridge University Press

- Koizumi, Tamotsu and Makino, Tutomu. 1971. "Oninron 1 (Phonology 1)" in *Eigo-gaku Taikei*, No. 1. Tokyo: Taishukan
- Kanehiro, Masao. 1978. "Eigo Onseigaku (The Phonetics of English)" Tokyo: Yamaguchi-shoten
- Kimura, Tuneo. 1968. "Eigo Hatuon-gaku Enshu (A Shorter Course in American English Pronunciation)" Tokyo: Kaitakusha
- Ladefoged, Peter. 1982. *A Course in Phonetics*. Harcourt Brace Jovanovich, Inc.
- . 1975. *Three Areas of Experimental Phonetics*. London: Oxford University Press
- Lado, Robert. 1964. *Language Teaching. A Scientific Approach*. McGraw-Hill
- Lado, Robert and Fries, Charles C. *English Pronunciation*. Ann Arbor / The University of Michigan Press
- Makino, Tutomu. 1977. "Eigo no Hatuon (English Pronunciation)." Tokyo: Tokyo-shoseki
- O'Connor, J.D. 1967. *Better English Pronunciation*. Cambridge University Press
- . 1974. *A Course of English Pronunciation*. London: The British Broadcasting Cooperation.
- Ota, Akira. 1965. "Nichi Eigo no Ontaikei no Hikaku (A Comparison of the Sound Systems of English and Japanese)." in *Gendai Eigo Kyoiku Koza*, No. 7. Tokyo: Kenkyusha
- Oguri, Keizo. 1969. "Eigo Onseigaku (English Phonetics)." Tokyo: Shinozaki Shorin
- Palmer, Harold E. and Blandford, F.G. 1964. *A Grammar of Spoken English*. W. Heffer & Sons Ltd.
- Pike, Kenneth L. 1975. *Phonemics*. Ann Arbor / The University of Michigan Press
- . 1976. *Phonetics*. Ann Arbor / The University of Michigan Press
- Quirk, Randolph. 1968. *The Use of English*. London: Longman
- Scott, Charles J. 1976. *Preliminaries to English Teaching*. Tokyo: ELEC
- Torii, Tugiyoshi and Kaneko, Naomichi. 1972. "Eigo no Hatuon (English pronunciation)" Tokyo: Taishukan
- Wijk, Axel. 1966. *Rules of Pronunciation for the English Language*. London: Oxford University Press
- "Eigo no Hatuon (English Pronunciation)" in *Gendai Eigo Kyoiku Koza*, No. 4. 1978. Tokyo. Kenkyusha
- "Onsei to Keitai." in *Nichi Eigo Hikaku Koza*, No. 1. 1985. Tokyo: Taishukan
- "Nichi Eigo no Hikaku (A Comparison of English and Japanese)" in *Gendai no Eigokyoiku* No. 8. Tokyo: Kenkyusha
- "Onseigaku Daijiten (A Dictionary of Phonetics)." 1976. Tokyo: Sanshusha
- Kenyon, John S. and Knott, Thomas A. 1953 *A Pronouncing Dictionary of American English*. Merriam.