Phonetic Analysis of Japanese University Students' Pronunciation of English Vowels

Martin Gore

(Received October 14, 1997)

1. Introduction

This paper presents a brief analysis of the phonetic problems encountered by Japanese university students in the pronunciation of English vowels and diphthongs. Unless otherwise stated, all remarks are made from the standpoint of British received pronunciation (RP), though, for the sake of wider applicability, observations that conflict with other major dialects of the English language have been avoided.

This study is based on the author's observations over the last twenty years teaching English to Japanese in London and in Japan, while also drawing upon the experience of phoneticians who have taught English to Japanese students, such as J.C.Catford, A.C.Gimson and Peter Ladefoged, through their published works.

Some graphical figures are presented to provide visual illumination of what is otherwise a conceptually confusing area, particularly in the discussion of the characteristics of English diphthongs. These figures were obtained using spectrograms taken in the phonetics class with a Kay Computerized Speech Laboratory, CSL 4300B (software version 5.X) and a Sure SM48 microphone.

2. Long and short vowels

/iː/

The English vowel /iː/ is so near to the corresponding Japanese vowel as to be indistinguishable from it in most contexts, and therefore presents comparatively few problems to students. The variation in vowel length between final fortis and lenis consonants (seat/seed) is important because the differentiation between these words is often achieved more by vowel length than by the quality of the final, the reduced form being in many situations approximately the same length as the short vowel /ɪ/ though retaining the tenseness of the closer vowel. See /ɪ/ below for more details.

The latter part of *twentieth*, which should have a clearly enunciated glide between two front close vowels with the second more open and back than the first, as /iɪ/, /iə/, /ie/ or /ɪə/,

is nearly always mistakenly pronounced with a long vowel /iː/ by the Japanese.

/ɪ/

Lower and laxer than the previous vowel, this vowel presents considerable problems to Japanese learners of English, who, as is pointed out under /ə/, tend to avoid the central area an in many cases perceive /ɪ/ as either /iː/ or /e/.

Both the qualitative and the quantitative distinctions between /iː/ and /ɪ/ must be observed. Thus, four types of vowel need to be practised: close tense long /iː/ (*seed*), close tense reduced /i/ (*seat*), half-close lax short /ɪ/ (*Sid*), half-close lax short reduced /ɪ/ (*sit*). Care must be taken to differentiate the last two from /e/ (*said,set*).

In his practical Course (1979), Gimson notes that reduced long vowels become so short that they may be of about the same length as short vowels, e.g. *feet* and *fit* are distinguished for practical purposes only by quality. Gimson (1980) quotes a study by Kalevi Wiik (1965) which, if applied to our example above, would give the following average vowel lengths in csecs: *seed* (36.0), *seat* (12.3), *Sid* (14.7), *sit* (7.3). This is a graphic instance of the phenomenon whereby "long" vowels before a fortis are typically as short as, or even shorter than, "short" vowels before a lenis.

/e/

This English vowel is, in isolation, similar to the Japanese corresponding vowel and seems to present comparatively few problems, but the reduction to /ɪ/ in unaccented syllables (as in the first vowel in *excuse me*) is not always correctly perceived, and the diphthong /eɪ/ (as in *A*) is often mistakenly pronounced as a long vowel /ɛː/ (see section on diphthongs, below).

/æ/

This is a difficult vowel for the Japanese, but one that most have mastered by the time they reach university. Confusions occasionally occur on the one hand with /e/ (bat-bet) and on the other hand with /ʌ/ (bat-but), but most Japanese students are aware of the importance of these oppositions and are able to make the distinctions clear, if not by totally accurate location, by use of constriction of the pharynx accompanied by the slight lengthening which tends to occur in the /æ/ syllable. There is, however, considerable confusion between *back* and *bag* resulting from transference from the corresponding Japanese words and from a failure to fully realize the importance of the extra lengthening before the voiced final consonant

in English. It would probably help here to use the notation /æː/ when the vowel is lengthened before a voiced final.

/ʌ/

It seems difficult for the Japanese to maintain a stable qualitative distinction between /ʌ/ and /ɑː/, and many tend to rely on the difference in length factor, which serves well on most occasions but not always. It is of utmost importance for Japanese students to have regular practice in the distinctions between /æ/, /ʌ/ and /ɑː/ as in *cat*, *cut*, *car*, without relying on length or /r/ retroflexion.

/ɑː/

This vowel poses few difficulties, but it must not be forgotten that it is more open than the corresponding Japanese vowel, which is closer to /a/.

/ɒ/

RP /ɒ/ is extremely close to the corresponding Japanese vowel (often notated /o/) and therefore causes few problems in itself. (There is some confusion, aggravated by English spelling and by the differences between British and American English pronunciation, over which words should have /ɒ/ , /aː/ , /əʊ/ , /ɔː/ and /ʌ/, as in cot, cart, coat, caught and cut, but this rarely affects the stability of /ɒ/.)

/ɔː/

As Gimson (1980) notes:

> The monophthongal nature of /ɔː/ should be insisted upon, especial care being taken to keep a proper distinction between /ɔː/ and /əʊ/ in such pairs as *caught*, *coat*; *saw*, *so*.

This is one of the most important distinctions, and one which is very often ignored in Japan. (See /əʊ/ in the section on diphthongs, below.)

/ʊ/

This short vowel is similar to the corresponding Japanese vowel, and so presents relatively little difficulty. The distinction from the reduced form of the back vowel /uː/ (*foot-boot*) is of course important, but the difficulty for the Japanese appears to be more with the backness and rounding of /uː/. A more important problem concerning /ʊ/ is that this pho-

neme tends to be used intrusively by Japanese learners of English in reduced syllables, or, as is common in Japanese words, after final consonants; places where a native English speaker might use /ə/, or, in the latter case, a slight aspiration (perhaps tainted with /ə/), or nothing at all.

/uː/

This back close vowel is usually accompanied by slight lip rounding. Japanese learners of English need to practise making a distinct locational difference between /uː/ and /ʊ/ and also need to make a special effort to add lip rounding which is uncharacteristic of their own language.

As with the distinctions between the full and reduced forms of /iː/ and /ɪ/, four types of vowel need to be practised: close fully long /uː/ (*food*), close reduced long /u/ (*boot*), half-close lax short /ʊ/ (*good*) and half-close lax reduced short /ʊ/ (*foot*). (See section on /ɪ/, above, for comparative vowel lengths.) It should be noted that the difference between *use* (v.) [juːz] and *use* (n.)[jus] is one of length reduction not vowel quality, though of course the devoicing of the final consonant is also important.

/ɜː/ and /ə/

The quality of these two central vowels is in most cases indistinguishable, so that /ə/ may be considered an "unaccented allophone" of /ɜː/ (Gimson, 1980).

As Gimson, and before him, D.B.Fry (1947), have pointed out, /ə/ is by far the most common vowel in colloquial English at 10.74% (total vowels being 39.21%) followed by /ɪ/ at 8.33% and /e/ at 2.97%. The diphthong /aɪ/ comes next at 1.83%.

This, as Gimson notes, is to be expected, since /ə/ is the most common vowel in unaccented syllables in a language which has a high proportion of unaccented syllables.

Consider the following (Gore, 1996):

I must go to see them has full forms: /aɪ mʌst gəʊ tʊ siː ðem/.

Reduction of unaccented syllables gives: /aɪ məs gəʊ tə siː ðəm/.

Owing to the high frequency of occurrence and the common role that /ə/ (and to a lesser extent /ɪ/) plays as a reduced-form substitute for other vowels including /ɜː/ in unaccented positions, the ability to recognize and produce a stable and natural /ə/, is a basic requirement of aural and oral competence. The difficulty Japanese learners face in accurately articulating these vowels is thus a major impediment to the acquisition of English pronunciation and to listening comprehension.

Apart from its common use within words, /ɜ:/ is also the sound most often used to indicate hesitation between words by native English language speakers. This, too, is difficult for the Japanese to master, and should be practised as an aim in itself.

Japanese students should aim to avoid dependence on /ɾ/ or similar colouring by retroflexion to secure a neutral /ɜ:/, even in cases where /ɜ:/ is spelt with a vowel + *r*.

The phonetic transcription of texts and recorded material is extremely helpful for university students in developing mastery over the occurrence, quality and quantity of central vowels. The use of computed spectrography is also helpful. /ə/ can be trained effectively using both the Sona-Match frequency-by-amplitude mode and also the vowel-chart mode. *Bird, girl, home* (/bə:d/, /gə:l/ and /həʊm/) and other such "schwa" words, which tend to become /ba:do/ , /ga:rʊ/ and /hɔ:mʊ/ etc., have been successfully trained by these (Gore, 1996)

3. Diphthongs

One problem common to nearly all the English diphthongs as pronounced by Japanese learners concerns the indeterminate and "gliding" nature of the English diphthong sounds.

Another, possibly related, problem is lack of confidence on the central vowels /ɜ:/ and /ə/.

We can see from a comparison of real-time formant frequency charts of diphthongs how the Japanese learner of English avoids the central vowel area as if fighting shy of indeterminate vowels (perhaps an apt verb because it may be in some part a question of courage -- the courage to utter with clarity a sound that is difficult to define and does not exist in one's native language). (See Gore, 1996.)

/eɪ/

One often hears Japanese university students pronounce the *A* of *A, B, C,* as if it were a long vowel /ɛ:/. This habit is reinforced by transference from Japanese, where /ɛ:/ is a permitted, and indeed standard, contraction of /ei/.

It is important that the learner should understand that English /eɪ/ is not a long vowel /e:/ or /ɛ:/ but a glide. The second element of the glide [ɪ], while being absolutely essential, should not be given separate pneumonic or other emphasis but be merely suggested, as if at the end of a continuous diminuendo which never quite reaches the region of the fully closed [i] and may indeed even stop short of a proper /ɪ/ position.

The same may be said of /aɪ/ and /ɔɪ/.

/aɪ/

Care should be taken not to confuse the English word *I*, which is a median glide /'aɪ/, and the Japanese noun of superficially similar value, *ai*, where both elements are peripheral and are given approximately equal emphasis (/'a-'i/). In computed spectrography, vowel-chart graphs of first formant against second formant of Japanese students' attempts to say *I* (/aɪ/) often show two small and widely separated areas, one (/a/) at F1: 720--880Hz and F2: 1,260--1,340Hz, and the other (/i/) at F1: 260--310Hz and F2: 2,700--2,800Hz (approx. 2,100Hz in men) with no transitional glide. (See Gore, 1996.)

It may be helpful for the Japanese to consider that the English *I* occurs in the gap between the Japanese *a* and *i*, without touching on either of these two extremes. Words such as *ear* and *here* present a similar problem in reverse. It is helpful to practise these "reverse" pairs together, especially if a spectrograph vowel chart is available to make real-time comparisons.

/ɔɪ/

As with the other English diphthongs, the gliding nature of this sound does not come naturally to the Japanese, and therefore needs practice. Again, care must be taken not to "overshoot" to /i/.

/əʊ/

Japanese learners often fail to recognize English diphthongs when pronounced with transitional glides by a native speaker, and perceive, for instance, *coat* as *court*, or (to give a more extreme but not uncommon example) *going* as *getting*.

The /əʊ/ glide should start fairly centrally and diminuendo in the direction of /ʊ/, accompanied by a gradual rounding of the lips. Many textbooks and dictionaries, especially those published in Japan, transcribe this sound as "ou." This would result in a glide in quite a different direction (Fig.1). Such a transcription may have been appropriate historically, but the starting point (and predominant element) of the glide is now clearly of a more central nature. This is particularly true of RP, but also holds true for most American speakers.

The inappropriate notation often found in Japan surely aggravates the difficulties in identifying and correctly reproducing the "o" sounds among Japanese learners of English.

/ɑʊ/

This is one of the few vowel sounds which cause little perceptible difficulty for Japanese learners. There seems to be, for the Japanese, an inherent ease in gliding from /ɑ / in the direction of /ʊ / without unduly emphasizing the second element, though sometimes one does hear an unnecessary emphasis here.

/ɪə/

Learners should be careful not to use a close (Japanese) /i/, but the half-close /ɪ/. It is quite difficult for the Japanese to emphasize this half-close vowel in the proper manner without it becoming /i/. However, the effort must be made. The use of the vowel-chart mode of the spectrogram is especially helpful here. See Figure 2.

Care must also be taken not to separate the elements too distinctly. The Japanese tend to pronounce *ear* as /'i-'a/ (the reverse of the problem noted in /aɪ/ -- again, it may be helpful to consider that the English *ear* occurs in the gap between Japanese *i* and *a*, without touching on either of these two extremes).

/ɛə /

Of all the English diphthongs featuring the central vowel, /ɛə / is the most problem-free from the point of view of vowel quality. This is probably because the two elements of this diphthong are so close to each other that the Japanese tend to perceive the glide as a long vowel /ɛ:/. This perception prevents the second element from receiving the heavy /-'a/ treatment that tends to befall the other centring diphthongs, and results in a pleasant (though often rather frontal) approximation of an RP /ɛə / glide. The difference in continuous speech is often imperceptible, and, as with the other centring diphthongs, the addition of an American or regional British post-vocalic *r* often virtually eliminates any vowel-quality differences that might otherwise have been problematic, but then of course there are often problems with the *r* itself. (These must be discussed at a later date.)

/ʊə/

This comparatively rare diphthong causes few difficulties, except that, as in the previous example, care should be taken to avoid /a/. It should be noted that in the speech of native English speakers, the glide in this diphthong is a slight one, often coalescing into /ɔə /or /ɔ:/, *your* becoming /jɔ:/ or, in unaccented situations, /jə/.

4. Closing remarks

It was the purpose of this paper to illuminate the problems that Japanese learners of spoken English have with English vowel sounds by a careful analysis of each separate item. The above comments have dealt only with short vowels, long vowels and diphthongs. Little reference has been made to other equally important areas of pronunciation, such as consonants or prosodic features such as stress groups in continuous speech. Diphthongs plus /ə/, and vowels in combination with semi-vowels and/or consonants present the same basic problems as delineated above but also bring new problems which will need to be analysed in conjunction with consonants. The difficulties the Japanese encounter in pronouncing /ji/ in *year*, /wʊ/ in *wood*, /si:/ in *seat*, initial and intervocalic /r/, final /l/ and other consonants are examples which may have bearing on vowel quality, but analysis of these problems is beyond the scope of the present paper.

Fig.1. Japanese student *so* superimposed on RP *so*.
The student's *so* starts at /o/ and ends at /u/, whereas RP *so* glides almost imperceptibly from /ə/ to /ʊ/. (Adult male voices.)

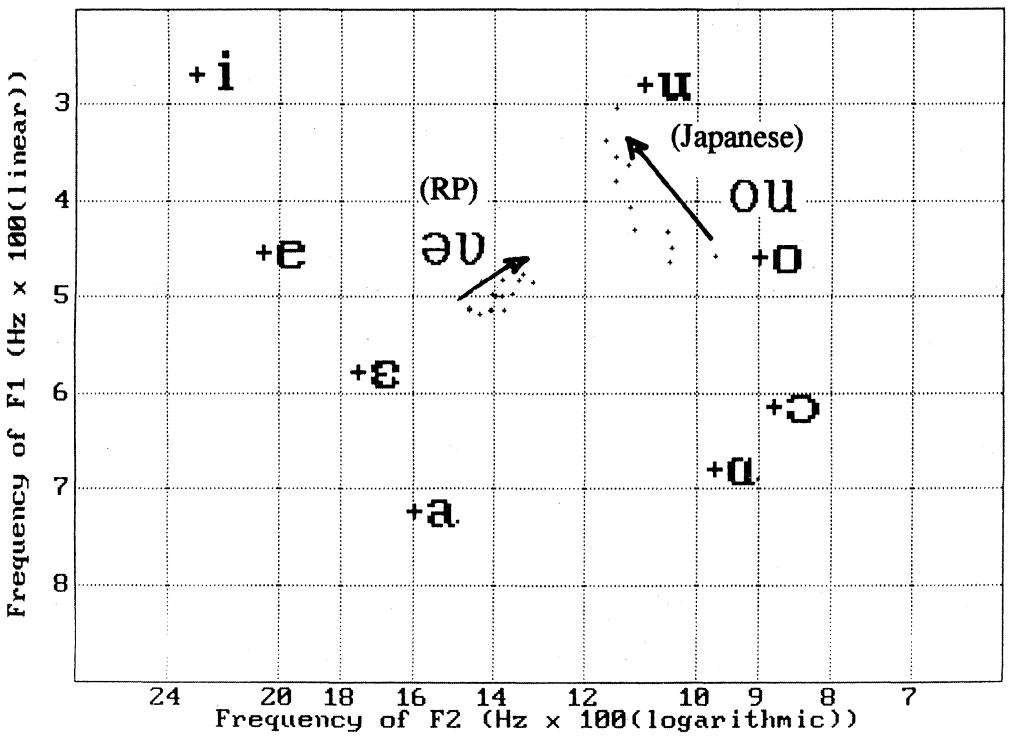
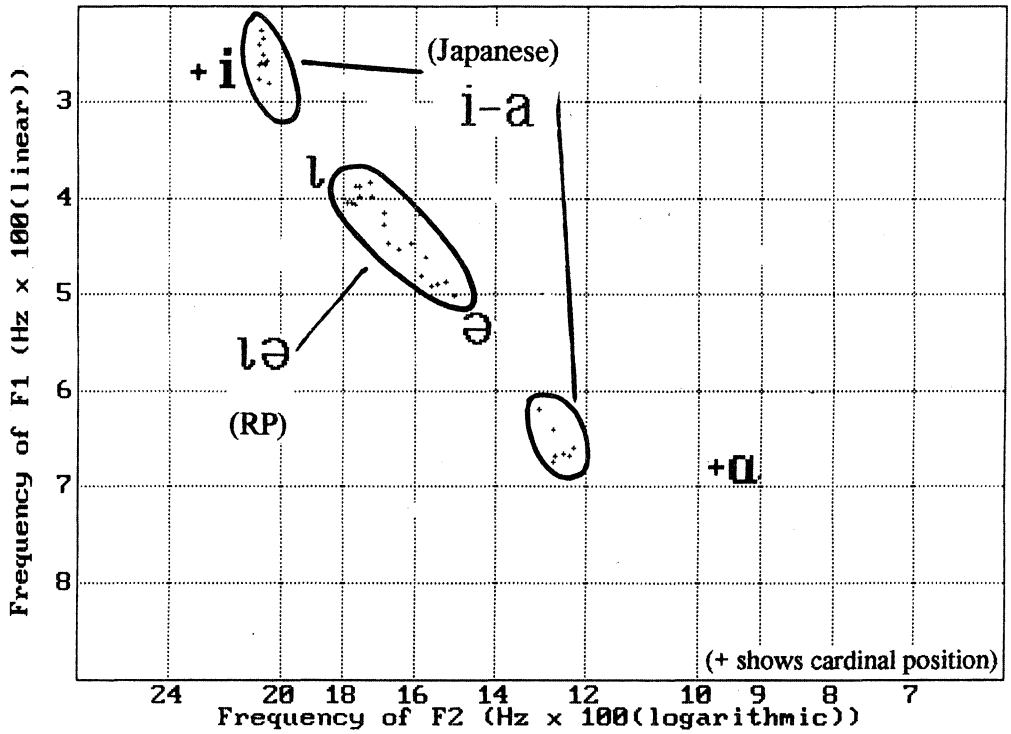


Fig.2. Japanese student *here* superimposed on RP *here*.
The student's *here* is split into two areas, /i/, and /a/,
whereas RP *here* glides from /ɪ/ to /ə/.



References

Catford, J. C. *A Practical Introduction to Phonetics.* Oxford, 1988.

Fry, D. B. "The Frequency of the Occurrence of Speech Sounds in Southern English." *Archives Néerlandaises de Phonétique Expérimentale, XX*, 1947.

Gimson, A. C. *Introduction to the Pronunciation of English.* --3rd ed., Arnold, 1980.

Gimson, A. C. *A Practical Course of English Pronunciation*: a perceptual approach. -- reprinted, Arnold. 1979.

Gore, M. *Formant Frequency Characteristics of English Central Vowels and Diphthongs Acquired by Japanese University Students.* Kagoshima Studies in English Language and Literature, 27. Kagoshima University, 1996.

Ladefoged, P. *A Course in Phonetics.* Harcourt Brace Jovanovich, 1975.

Wells, A. C. *Pronunciation Dictionary.* Longman, 1990.