ACOUSTIC INVESTIGATION OF ENGLISH AND JAPANESE [s] AND [θ] BY ENGLISH AND JAPANESE SPEAKERS

Miho Kamata

Abstract

It has been said that there are certain differences between English /s/ and Japanese /s/ in their articulation, and also in their auditory impression that the former sounds stronger than the latter.

This paper is trying to reveal the differences/similarities of the acoustic characteristics of English and Japanese voiceless alveolar fricatives in their L1 and L2 speech and to know how the articulatory differences and the auditory impression would be reflected in acoustical properties such as their amplitude, duration, formant frequency, and spectral shape.

Three native British English speakers whose L2 is Japanese, and three native Japanese speakers whose L2 is English were asked to read out word lists in their first and second languages. As a result, it was found that there was an interesting prominence in the lower frequency region of the spectral shape that can be seen only in the spectra of alveolar fricatives uttered by the English speakers but not in those uttered by the Japanese speakers.

1. Introduction

A voiceless alveolar fricative, /s/, is reported by Nartey (1982) to be the most frequently occurred sound among the voiceless fricatives which can be heard in 317 languages in the world (Kent and Read 1992: 129).

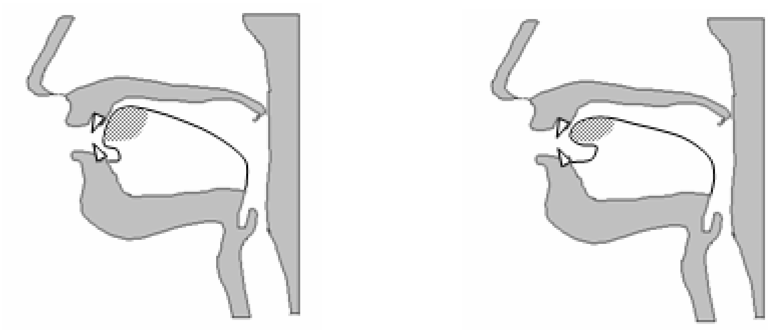
Also in English and Japanese, which we will deal with in this paper, the voiceless alveolar fricative is commonly used and it has an established phoneme in the phonemic inventories of each language. It seems, however, that their realisations are slightly different depending on speakers of each language. In this paper, we will investigate voiceless alveolar fricatives elicited from six native speakers of English and Japanese reading both English and Japanese word lists as their L1 (native language) and L2 (second language) in order to know their acoustical similarities and/or differences not only in their L1 and L2 speech but also in the L1 English and the L1 Japanese.

2. Literature Review

My interest in this topic stems from two sources. One is my intuitive impression that /s/ in English sounds stronger and more fricated than that in Japanese; this impression is reinforced by another source, which is the notion of some scholars, such as Ono (1986: 68) and Takebayashi (1996: 40), saying that English /s/ is somewhat stronger than Japanese /s/. In the next section we will overview the notion of these scholars in the past, and then, we will have a look at phonemic/phonetic and acoustical characteristics of the sounds concerned in both languages in the rest of this section before we analyse and discuss them in later sections.

2-1. English /s/ vs. Japanese /s/ — auditory & articulatory difference

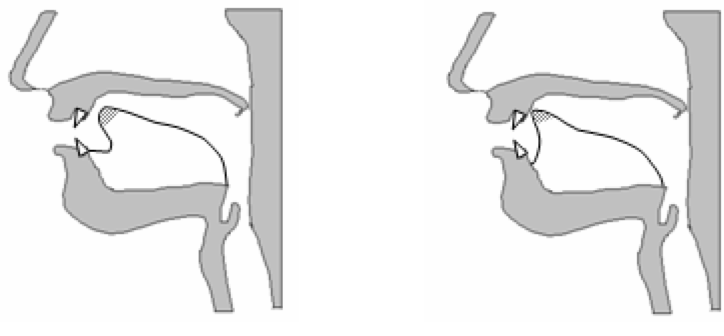
Some Japanese phoneticians have pointed out that English /s/ is stronger than Japanese /s/ (Ono 1986: 68). Takebayashi (1996: 40) also points out that frication of English /s/ is stronger than that of Japanese /s/; he explains this difference is normally due to the difference of the place of articulation in both languages. /s/ uttered by both native speakers of English and those of Japanese are described generally as *voiceless alveolar fricatives* and specifically as *lamino-alveolar fricatives* (*ibid*.), which are articulated with the blade of the tongue and the alveolar ridge forming the close approximation, in both languages, but strictly speaking, they are slightly different in terms of tongue position in their articulation. According to Saito and Takebayashi (1998: 97), lamino-alveolar fricatives spoken by English and American native speakers of English tend to be articulated with the tip of the tongue raised towards the alveolar ridge as shown in Figure1-(B), while those spoken by Japanese tend to be articulated with the tip of the tongue lowered close to the lower front teeth as shown in Figure1-(A):



(A) **(B)**

Figure 1. The place of articulation for lamino-alveolar fricatives
(Based on Takebayashi 1996: 40)

In addition, according to Takebayashi (1996: 221), those who use *apico-alveolar fricatives* are recently increasing among young, especially female, speakers of Japanese. Those speakers articulate the sound with less friction noise with the tip of the tongue raised towards the alveolar ridge (Figure2-(B)) or towards the front part of the alveolar ridge, closer to the teeth (Figure2-(A)). In this case, the resultant sound will be more like dentalised /s/ (i.e. [s̪]), or simply become [θ] (Figure3).



(A) **(B)**

Figure 2. The place of articulation for apico-alveolar fricatives

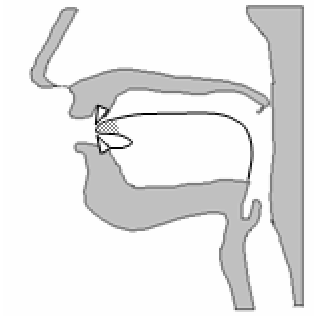


Figure 3. The place of articulation for dental fricatives
(Based on Takebayashi 1996: 200)

Takebayashi (1996: 201) also points out the difference between English /s/ and Japanese /s/ in their posture of the tongue; the surface of the tongue tip or blade for Japanese /s/ is flatter (like the one for /θ/ in Figure 4) than the one for typical English /s/ which is deeply grooved in general (as shown in Figure 4), so that the frication of Japanese /s/ sounds weaker than the one of English /s/.

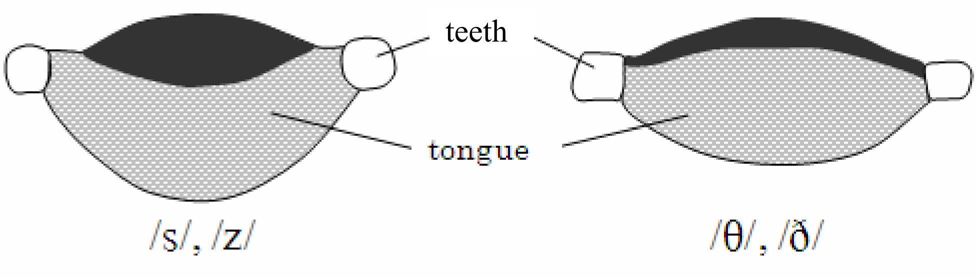


Figure 4. The tongue posture of /s, z/ and /θ, ð/ (front view)
(Based on Takebayashi 1996: 201)

2-2. General Characteristics – phonemic & phonetic status of [s] and [θ] in English and Japanese

Given the difference of voiceless alveolar fricatives in English and Japanese, the main realisations of /s/ in both languages can be listed in Table 1 below:

Table 1. English /s/ vs. Japanese /s/

		Place of articulation			Realisations
English /s/	–	lamino-alveolar fricative	(Figure1-(B))	→	[s̻]
Japanese /s/	–	lamino-alveolar fricative	(Figure1-(A))	→	[s̻]
		apico-alveolar fricative	(Figure2-(B))	→	[s̺]
		dentalised alveolar fricative	(Figure2-(A))	→	[s̪]
		or dental fricative	(Figure3)	→	[θ]

/s/ in English does actually vary dialectally and from an interspeaker perspective.

Stuart-Smith (2004) cites, for example, that Macafee (1983: 34) has noted an apico-alveolar articulation in working-class speech in Glasgow, which also occurs more generally in urban varieties of Scots (Johnston 1997: 509); this articulation sounds auditorily 'retracted', and is thought to be produced with the tongue tip raised, thus opening up the sublingual cavity, lengthening the front-cavity and lowering the resonance so that it sounds rather like /ʃ/. She also found this distinctive 'retracted' pronunciation of /s/ as a property of male, especially boys', speech in her own data (Stuart-Smith 2004). Regarding interspeaker variation, we can observe enormous interspeaker variation in how dentalised/apicalised an /s/ is, and, at the other extreme, how it tends towards /ʃ/. There is also an intraspeaker variation in rounding which seems very common. However, we do not go into the detail of these dialectal or inter/intraspeaker variation in English in this paper. Accordingly, it safe to say that English speakers *mainly* pronounce /s/ with lamino-alveolar fricative [s̻] as its principal allophone as shown in Table 1. In the case of Japanese, Japanese speakers mainly pronounce /s/ with three or four possibilities ([s̻], [s̪], [s̺] or [θ]) depending on their tongue position and shape.

It should be noted that the substitution of a dental fricative [θ] for an alveolar fricative [s] does not usually occur among English speakers, because in English these two sounds are phonemically different; in other words, English has two different phonemes, an alveolar fricative /s/ and a dental fricative /θ/. Therefore, in English, if the first consonant in 'sigh' were pronounced with a dental realisation [θ] instead of an alveolar one [s], the lexical meaning of the word would be totally changed to 'thigh'. This also reflects the fact that [s, z] are made with a configuration of the tongue which allows the air to escape along a groove (c.f. grooved aspect), whereas [θ, ð] use a flatter configuration where the air escapes through a slit (c.f. slit aspect) (Laver 1994: 140, 258-260, Cruttenden 2001: 177-178).

In the case of Japanese, however, these two sounds are not phonemically distinguished; that is to say, Japanese has only one phoneme /s/ whose possible place of articulation normally covers the region of both places of articulation of English /s/ and /θ/. In other words, Japanese does not have a dental fricative phoneme, but a dental allophone of /s/ could occur. Therefore both [saɪ] and [θaɪ] could be pronunciations of the same word /saɪ/ for certain Japanese people who are not sensitive to other languages which distinguish these two sounds phonemically, such as English.

This fact is reinforced by the perceptual map of fricatives (Figure5) made by Johnson (2003: 132-133) from the confusion matrices published by Lambacher, Nelson and Berman (2001). It shows the results of two multidimensional scaling analyses – one for American English listeners (black letters on white background) and one for Japanese listeners (white letters on dark background). All of the listeners heard the same American English stimuli, and identified them as "f", "th", "s", "sh", or "h". The most significant point from the figure for the present study is that [θ] is closer to [s] than to any other fricative in the Japanese perceptual space. Given this result, Johnson presumes an interesting assumption for it in addition to pointing out the fact that [θ] is not a surface allophone in Japanese. He says that though the [s] and [θ] frication noises are quite different from each other, their formant transitions (an important cue for place of articulation) are quite similar; therefore, the Japanese listeners may be relying on this similarity and ignoring the frication noise differences that seem so significant for the American English listeners.

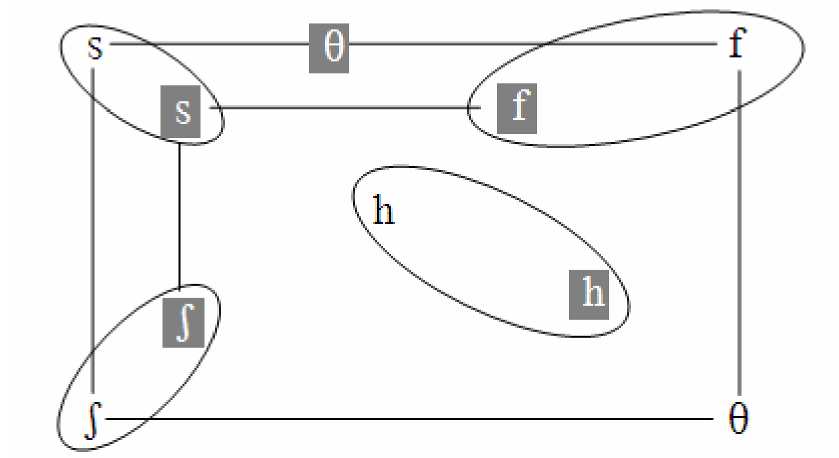


Figure 5. Multidimensional scaling analysis of data reported by Lambacher *et al*. (2001): the perceptual map of fricatives for American English listeners (black letters on white background) and Japanese listeners (white letters on a dark background) (Based on Johnson, 2003: 132).

Table 2 shows the phonemic differences between English and Japanese in terms of place of articulation:

Table 2. Coronal and anterior voiceless fricatives in English and Japanese

	Place of articulation	
	dental	alveolar
English	/θ/	/s/
Japanese	/s/	

This absence in Japanese of dental fricative phonemes and the perceptual similarity of the dental fricatives to the alveolar fricatives allow a dentalised quality in the alveolar articulations, which is liable in English to cause confusion with /θ/ or to produce a *lisping* fricative which is considered socially undesirable. Therefore this kind of lisping, i.e. the substitution of /θ/ for /s/ or strongly dentalised version of /s/ is considered as a common speech defect in English (Cruttenden 2001: 178, Takebayashi 1996: 201).

In this way it has been said that English /s/ is stronger than Japanese /s/ and it seems very likely that the weaker Japanese /s/ is in fact dentalised or realised with dental fricatives. However, how can these subjective impressions such as ‘stronger’ or ‘weaker’ be shown in the objective acoustical characteristics such as in a spectrogram or spectrum? In order to know, we will look at the general acoustic characteristics of fricatives and some experimental studies and findings in the next section.

2-3. Acoustic Characteristics of /s/ and /θ/

Here we will look at the main acoustic characteristics (intensity, duration and frication noise frequency) not only of /s/ but also of /θ/ because of the fact that some Japanese substitute [θ] for /s/ in Japanese.

2-3-1. Intensity

It is generally said that alveolar fricatives are more intense than dental fricatives. According to Cruttenden (2001: 180) and Hayward (2000:190), relative intensity is particularly important for distinguishing between sibilant and non-sibilant fricatives; in English, for example, as far as voiceless fricatives are concerned, sibilant /s/ and /ʃ/ are clearly more intense than non-sibilant /f/ and /θ/.

How can we measure intensity then? Intensity is the amount of energy carried by a sound wave, measured in *decibels* (Trask 1996: 181); the perceptual correlate of intensity is *loudness*, although the relationship is far from linear. Also the characteristics of the *amplitude* in a sound wave most directly relate to the intensity of the associated sound and hence, in a complex manner, to the perceived loudness (Trask 1996: 21). Therefore it would be considered to be more adequate to take the measurements of the amplitude of the waveform we are concerned with.

According to Johnson (2003: 31), the amplitude of a waveform can be measured in several ways. In simple periodic (sine) waves, the three methods – peak amplitude, peak-to-peak amplitude, and root mean square (RMS) amplitude – can be mathematically derived from each other[1]. For complex waves, however, the different measures of amplitude are not mutually predictable. The peak measurements differ from RMS amplitude in that they give a measure of acoustic amplitude, whereas RMS amplitude is a measure of acoustic intensity, i.e. the sum of all frequency amplitudes[2].

2-3-2. Duration

Fricatives as well as stops and affricates involve noise generation. It is risky to assign a particular duration to fricative noise segments because the duration is influenced by numerous contextual factors. However, compared to the stops and affricates, fricatives have relatively long durations of noise of aperiodic energy. (Kent and Read 1992: 121).

According to Kent and Read (1992: 122), You (1979) found that the duration of noise for a fricative varies with place of articulation with average duration increasing in the following order: dentals, labials, alveolars, and palatals. Therefore, if this is always true, the duration of a voiceless alveolar fricative [s] should be longer than that

[1] In simple periodic (sine) waves, the peak-to-peak amplitude is twice the peak amplitude, and RMS amplitude taken over one cycle of a sine wave is equal to the peak amplitude multiplied by 0.707 (Johnson 2003: 31).

[2] In order to calculate RMS amplitude, each sample in a waveform window is squared; then the average of the squares is calculated; and finally, the square root is taken: i.e. the root of the mean of the squared samples (Johnson 2003: 31-33).

of a voiceless alveolar fricative [θ] or a voiceless dentalised-alveolar fricative realisation [s̪] in equivalent contexts.

2-3-3. Frication noise frequency

2-3-3-1. Noise frequency

According to Kent and Read (1992: 123), the alveolar fricative has more energy at higher frequencies compared to the palatal; as a rule of thumb for adult male talkers, the major region of noise energy for the alveolar fricatives lies above 4,000 Hz, while the palatal fricatives have significant noise energy extending down to about 3,000 Hz. A perceptual test of synthesised fricatives by Heinz and Stevens (1961) modelling these sounds with a single low-frequency zero (antiresonance) and a single pole (resonance) applied to a white noise source revealed that listeners identified the resulting noise as /s/ when the centre frequency of the pole was between about 4,000-8,000 Hz and as /ʃ/ when the centre frequency was below about 3,000 Hz (Kent and Read 1992: 123-124).

Similarly, according to Cruttenden (2001: 180), continuous noise in the spectrum is appropriate to articulatory friction regions; an alveolar fricative has continuous noise at a range from 3,600 Hz to 8000 Hz in the spectrum, while a dental fricative at a range from 1,400 Hz to 8000 Hz in the spectrum. As we will see in the following section (§2-3-3-3), the former has a “compact” spectrum and the latter has a “diffuse” spectrum.

2-3-3-2. Frequency region before the following vowel’s F5

According to Steven’s perceptual experiments in 1985, the frequency region around a vowel’s F5 may be crucial for distinguishing between [s] and [θ]. In his experiments using synthetic sounds, he found that if the frication noise for the fricative was stronger than the vowel in the F5 region, listeners heard it as [s], and if the friction noise was weaker than the vowel in this region, listeners heard as [θ].

For his [θ]-[s] identification test, according to Stevens (1985: 244-246), a series of eight fricative-vowel syllables (with the vowel [ɑ]) was synthesised in such a way that the amplitude of the turbulence noise at high frequencies in the F5-F6 region was gradually reduced from one stimulus to the next in the series. The change in spectrum amplitude in this region from one stimulus to the next was 5dB in a range of –18dB to +17dB. The first three formants underwent transitions during the initial 30-40 msec following the onset of voicing. Then a number of replications of these synthetic syllables was presented in random order to four American English listeners, who identified the consonant as either [θ] or [s]. The results show that the listeners appear to identify the consonant as [θ] if the high frequency noise amplitude increases at the fricative-vowel boundary and as [s] if there is a drop in amplitude as the frication noise ends and voicing for the vowel begins, with only one stimulus (i.e. –3dB) that was not identified unanimously as [s] or [θ] by all listeners.

2-3-3-3. Shape of spectra – compact vs. diffuse

It should first be noted here that, as Hayward (2000: 191) mentions, the spectra of

English fricatives vary considerably from speaker to speaker; that is, while the observations just made are generally valid, it would be inappropriate to add more detail. At least for English, however, according to Hayward (2000: 191), it has seemed appropriate to describe fricative spectra in terms of the general categories we might apply to the spectra of stop bursts (diffuse vs. compact, high-frequency prominent vs. low-frequency prominent) rather than in terms of specific formant frequencies, as would be the case for vowel spectra. Now let us look at Figure 6 below, which shows average spectra for the English voiceless fricatives /θ s/ extracted from the words *thigh* and *sigh*, pronounced by an English male speaker:

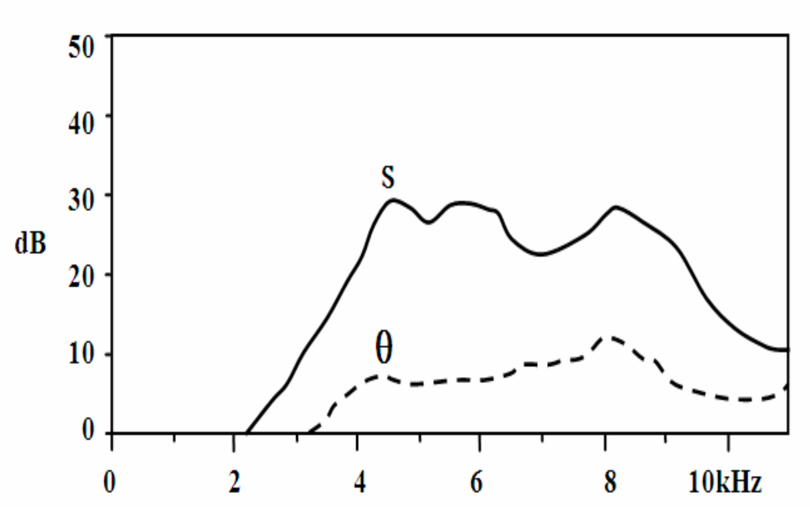
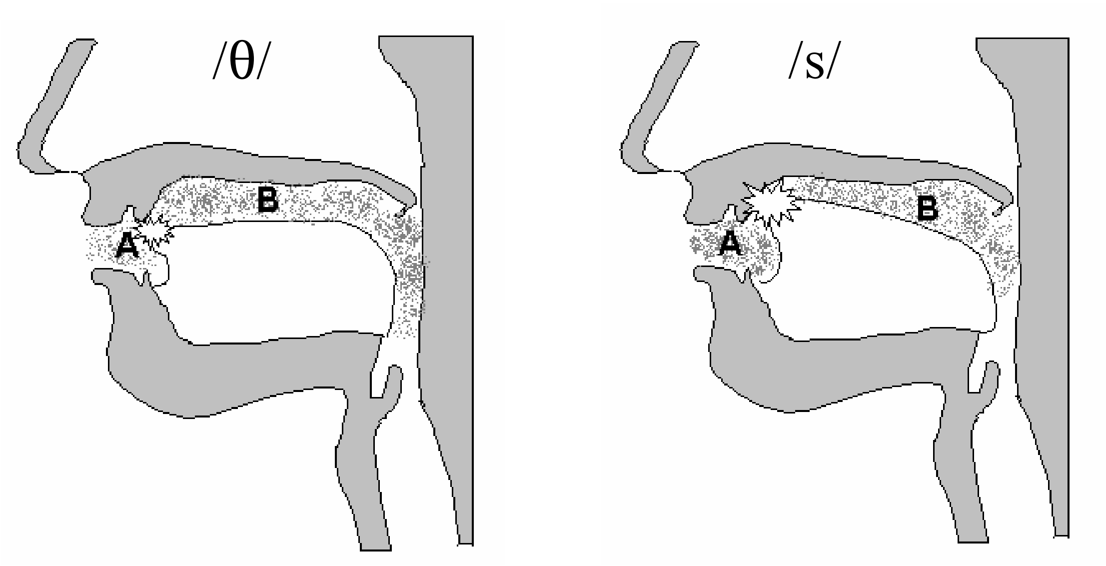


Figure 6. Average spectra for the English voiceless fricatives /θ s / extracted from the words *thigh* and *sigh* pronounced by an English male speaker
(After Hayward 2000: 191)

Figure 6 tells us an interesting feature which distinguishes /s/ and /θ/: that is, the difference of the shape of spectra. The spectrum of /s/ is more compact, while that of /θ/ is more diffuse in character; to be concrete, in the *compact* spectrum of /s/, on the one hand, a concentration of energy can be seen in the middle of the frequency axis, in the *diffuse* spectrum of /θ/, on the other hand, the energy is not so concentrated in a particular frequency region. This distinction was introduced by Jakobson, Fant and Halle (1952). According to Jakobson *et al.* (1952:27), the essential articulatory difference between the compact and diffuse phonemes lies in the relation between the volume of the resonating cavities in front of the narrowest stricture (**A** in Figure 7) and those behind this stricture (**B** in Figure 7); the ratio of the former to the latter (**A**:**B**, i.e. **A**/**B**) is higher for the compact than for the corresponding diffuse phonemes.



A/B=lower → diffuse spectrum A/B=higher → compact spectrum

⇒the ratio of A to B is higher for /s/ whose spectrum is compact, and lower for /θ/ whose spectrum is diffuse.

Figure 7. The volume difference of the resonating cavities in front of the narrowest stricture (A) and those behind it (B)

Since the point at which the constriction is made for /θ/ is more forward than that for /s/ as can be seen in Figure 7, the ratio of A to B for /θ/ is lower, resulting in the diffuse spectrum, and that for /s/ is higher, resulting in the compact spectrum.

Johnson (2003: 124-127) also explains that the frequency change in fricatives is associated with the place of articulation; the frequency change is caused by changes in the filtering action of the vocal tract, and in particular, changes in the length of the front cavity of the vocal tract, which we have seen as **A** in Figure 7. To be concrete, the shorter the front cavity (i.e. **A** in Figure 7), the higher the frequency of the lowest spectral peak (except for the case with no front cavity). According to Kent and Read (1992: 35), when the front cavity is very short, its lowest resonance frequency is too high to offer appreciable shaping of the noise energy. Consequently, the spectrum for these fricatives is flat or diffuse, lacking prominent peaks or valleys.

3. Research Topics and Expectations

Based on these articulatory and acoustic characteristics of the sounds that we are concerned with, we would like to know whether the articulatory or impressionistic notions of some scholars and myself to the difference between English /s/ and Japanese /s/, which are mentioned in the earlier parts in this paper, are acoustically true or not. The following research topics will be considered in this paper.

First, we will consider the differences and/or similarities between English /s/ and Japanese /s/ by their L1 speakers. As it seems that English /s/ is stronger than Japanese /s/, we can expect that the overall amplitude of English /s/ is higher than Japanese and/or duration of English /s/ is longer than Japanese.

Secondly we will consider /s/ in the English context and Japanese context respectively. As for /s/ in the English context, English /s/ uttered by its L1 speakers (i.e. the English) and L2 speakers (i.e. the Japanese) will be compared. Similarly, as for /s/ in the Japanese context, Japanese /s/ by its L2 speakers (i.e. the English) and L1 speakers (i.e. the Japanese) will be compared. In both cases, we can expect that the former /s/ (i.e. /s/ produced by English speakers) shows higher amplitude, longer duration than the latter /s/ (i.e. /s/ by Japanese speakers) does.

Thirdly we will also compare English /θ/ by its L1 (i.e. English) speakers with that by its L2 (i.e. Japanese) speakers. Given the fact that there is not a dental fricative phoneme in Japanese and given the perceptual map of fricatives, reported by Johnson (2003: 132-3), which shows /θ/ is closer to /s/, we can predict that /θ/ is highly possibly pronounced more like [s] by Japanese speakers; if so, their /θ/ will be expected to show higher amplitude at their noise frequency peak and longer duration than /θ/ by the English speakers.

4. Methods

4-1. Subjects

Three female native speakers of (British) English, who speak Japanese as their second language, and three female native speakers of Japanese, who speak English as their second language and whose model of English is British rather than American will be used as subjects. Given the literature saying that Japanese young female speakers seem to pronounce more dentalised fricatives for alveolar fricatives (Takebayashi 1996: 221), all the subjects were female young speakers, aged from 20 to 24. Information about the speakers is listed in Table 3 below:

Table 3. Speakers' Information

ID	Nationality	Sex	Age	Birth Place and Home Town	Year of Study L2 / Year of Stay in Japan or England
E01	British	F	21	Bradford, England	4 yrs / 1yr in Tokyo, Japan
E02	British	F	22	Manchester,England	4 yrs / 1yr in Fukuoka, Japan
E03	British	F	23	Norwich, England	7 yrs / 10 mths in Fukuoka, Japan
J01	Japanese	F	22	Fukuoka, Japan	4 yrs / 3 yrs in Leeds, England
J02	Japanese	F	24	Miyazaki, Japan	12 yrs / 2 yrs in Leeds, England
J03	Japanese	F	20	Tokyo, Japan	8yrs/8yrs in London&Leeds,England

4-2. Data elicitation

The speech data were all elicited through word lists. All the speakers were asked to read Japanese and English word lists. The word lists were created in order to include words which provide the target sounds /s/ only in the word initial position, because in Japanese, /s/ occurs only in the syllable-initial (i.e. onset) position. This is because Japanese is a CV(N) language and consonants do not occur in the syllable-final (i.e. coda) position unless they are nasals (e.g. *ha<u>n</u>da<u>n</u>* /handaɴ/ 'judgement') or parts of double consonants (e.g. *ki<u>t</u>te* /kitte/ 'stamp'). Both English and Japanese word lists are listed in Table 4:

Table 4. Word lists and their phonemic transcriptions

	English word list		Japanese word list[3]
1e	Say 'say' to me. /seɪ seɪ tuː miː/	1j	kore wa tiisai 'sei' to iimasu. /kore wa tɕiːsai seː to iːmasɯ/ (trans.) This is a small 'surname'.
2e	Say 'sigh' to me. /seɪ saɪ tuː miː/	2j	kore wa tiisai 'sai' to iimasu. /kore wa tɕiːsai sai to iːmasɯ/ (trans.) This is a small 'difference'.
3e	Say 'sough' to me. /seɪ saʊ tuː miː/	3j	kore wa tiisai 'sau' to iimasu. /kore wa tɕiːsai saɯ to iːmasɯ/ (trans.) This is a small 'right&left'.
4e	Say 'saw' to me. /seɪ sɔː tuː miː/	4j	kore wa tiisai 'sou' to iimasu. /kore wa tɕiːsai soː to iːmasɯ/ (trans.) This is a small 'layer'.
5e	Say 'soy' to me. /seɪ sɔɪ tuː miː/	5j	kore wa tiisai 'soi' to iimasu. /kore wa tɕiːsai soi to iːmasɯ/ (trans.) This is a small 'plain clothes'.
6e	Say 'Sue' to me. /seɪ suː tuː miː/	6j	kore wa tiisai 'suu' to iimasu. /kore wa tɕiːsai sɯː to iːmasɯ/ (trans.) This is a small 'breathe'.
7e	Say 'thay' to me. /seɪ θeɪ tuː miː/		
8e	Say 'thigh' to me. /seɪ θaɪ tuː miː/		
9e	Say 'thou' to me. /seɪ θaʊ tuː miː/ (see note 7)		
10e	Say 'thaw' to me. /seɪ θɔː tuː miː/		
11e	Say 'thoy' to me. /seɪ θɔɪ tuː miː/		
12e	Say 'thew' to me. /seɪ θuː tuː miː/		

Points to notice when the lists were elaborated should be mentioned here. Firstly all of the words including target sounds were embedded in carrier phrases, such as "SAY ____ TO ME" in English, and "KORE WA TIISAI ____ TO IIMASU (これは小さい____と言います)" (*This is a small*____) in Japanese, in order for them to occur between vowels. The reason for this is because it is easier to measure the duration of frication noise of fricatives when it occurs between vowels.

The second point is that the contexts in which the target fricative sounds appear in English was made to correspond to those in Japanese, adopting /eɪ/ for English and

[3] The actual Japanese word list was written in Japanese with the reading of a Chinese character by writing kana (i.e. the Japanese syllabary) above it, not as it is written here in the Romanization system.

/ai/ for Japanese as the preceding vowels, and /eɪ, aɪ, aʊ, ɔː, ɔɪ, uː/ for English and /eː, ai, aɯ, oː, oi, ɯː/[4] for Japanese as the following vowels[5]. These vowels in English and Japanese can be considered fairly similar to each other, as Figures 8 and 9 show:

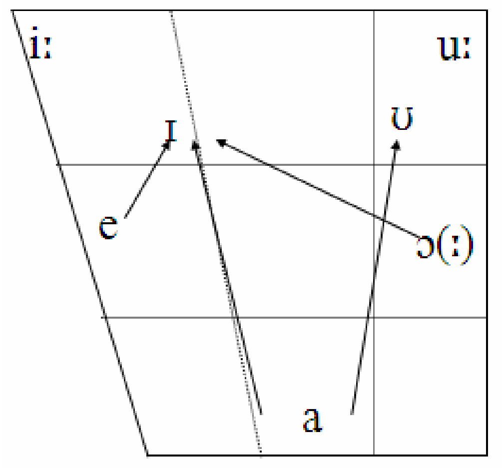


Figure 8. English Vowels concerned

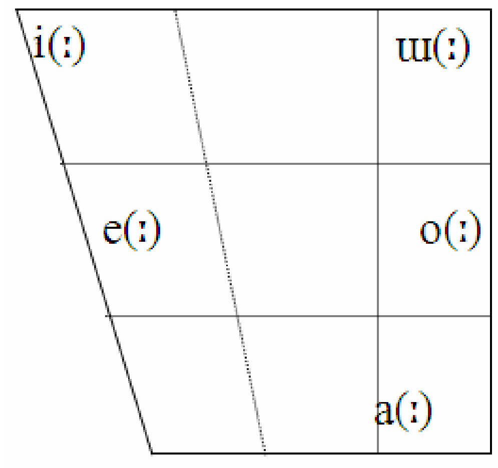


Figure 9. Japanese Vowels

One thing to be noted here is, although Japanese has these five vowels as shown in Figure 9, one of them, the high front vowel /i/, does not occur after a voiceless alveolar fricative in Japanese[6]. That is why the other four vowels /e a o ɯ/ were chosen as the sounds following /s/, and in English, /e a ɔ u/ were chosen as their corresponding vowels.

Thirdly the English word list has six extra sentences including the voiceless dental fricative /θ/ in the target words, while the Japanese one does not; this is only because the dental fricative is not a phoneme in Japanese.

Finally both the English and Japanese word lists include some non-sense words or uncommon words. In the case of the English word list, the words 'thay' /θeɪ/ (7e), 'thou' /θaʊ/[7] (9e) and 'thoy' /θɔɪ/ (11e) are all non-sense words; they are only coined to represent the sounds corresponding to English 'say' /seɪ/ (1e), 'sough' /saʊ/ (3e) and 'soy' /sɔɪ/ (5e) and Japanese 'sei' /seː/ (1j), 'sau' /saɯ/ (3j) and 'soi' /soi/ (5j). In the case of the Japanese word list, the pronunciation /saɯ/ of 'sau' (3j) is not very common.

---

[4] Japanese has five 'short' vowels, /i/, /e/, /a/, /o/ and /ɯ/ as shown in Figure 9; in addition, it also has five sets of succession of identical vowels, /iː/, /eː/, /aː/, /oː/ and /ɯː/, usually called 'long' vowels. Moreover, Japanese people generally do not have a notion of 'diphthong' in their language; instead, they treat the vowels such as [ai] or [oi] as a sequence of two sounds as if they are /a/+/i/, or /o/+/i/.

[5] In Japanese, the sequence of two sounds /e/+/i/ is actually pronounced as [eː] instead of [ei] which could be phonetically closer to English /eɪ/; consequently the sequence of /a/+/i/ and /eː/ were selected as the corresponding sounds to English /eɪ/ alternatively.

[6] The high front vowel /i/ can follow /s/ phonemically as in *sika* /sika/ 'deer'; however, it is phonetically realised not as [sika] but as [ɕika].

[7] The word 'thou' /ðaʊ/ is not a non-sense word; however, the word 'thou' with a pronunciation /θaʊ/ is a non-sense word.

4-3. Procedure

All the speakers were asked to read out all the sentences in both the English and Japanese word lists – reading each sentence once, and without any pauses within each sentence but with pauses between sentences. However, some speakers could not help putting a pause within a sentence, especially immediately before the target sound was pronounced, not only because of the structure of the sentence which could allow them to put a pause there, but also, probably, because of some unfamiliar and non-sense words in the lists. They were also asked to repeat the same sentence if they misread it.

All the speakers were asked to read the word list of their L2 language first; so they could not predict the L2 list from having read the L1 list.

4-4. Data Analysis

The way of analysing the data is using instrumental analysis. The following three acoustic measurements are taken; amplitude, duration and formant frequency.

As for amplitude, not only the amplitude of noise frequency peak, but also the average amplitude of the following vowel were measured in order to standardise it, calculating the ratio of the noise amplitude to vowel amplitude.

As to duration, the duration of noise friction of the fricatives as well as the following vowel were measured in order again to standardise them by calculating the proportion of the fricative duration in CV duration.

With reference to the formant frequency, the frequency of the noise peak (the lowest spectral peak), and that of the region before the following vowel's F5 were measured.

5. Results

Table 5 below shows all measurements of acoustic components (1)-(11). All the measured values for them as well as the phonetic realisations of all the tokens are shown in Appendix 1. Contrary to our expectations, no tokens of /s/ uttered by Japanese speakers were auditorily full dental fricatives [θ], although some English dental fricative phoneme /θ/ were realised as [s].

Table 5. Measurements of acoustic components (1)-(10) for /s/ and /θ/

No.	Acoustic Components
(1)	Frequency peak of the average spectrum of fricatives (Hz)
(2)	Amplitude of (1) (dB)
(3)	Frequency peak of the average spectrum of the vowels following fricatives (Hz)
(4)	Amplitude of (3) (dB)
(5)	Duration of fricatives (ms)
(6)	Duration of the following vowels (ms)
(7)	F5 of the following vowels (first 50 ms) (Hz)
(8)	Average Amplitude of the frequency region of (7) of fricatives (last 50 ms) (dB)
(9)	Average Amplitude of (7) (dB)
(10)	difference between (8) and (9) (dB)

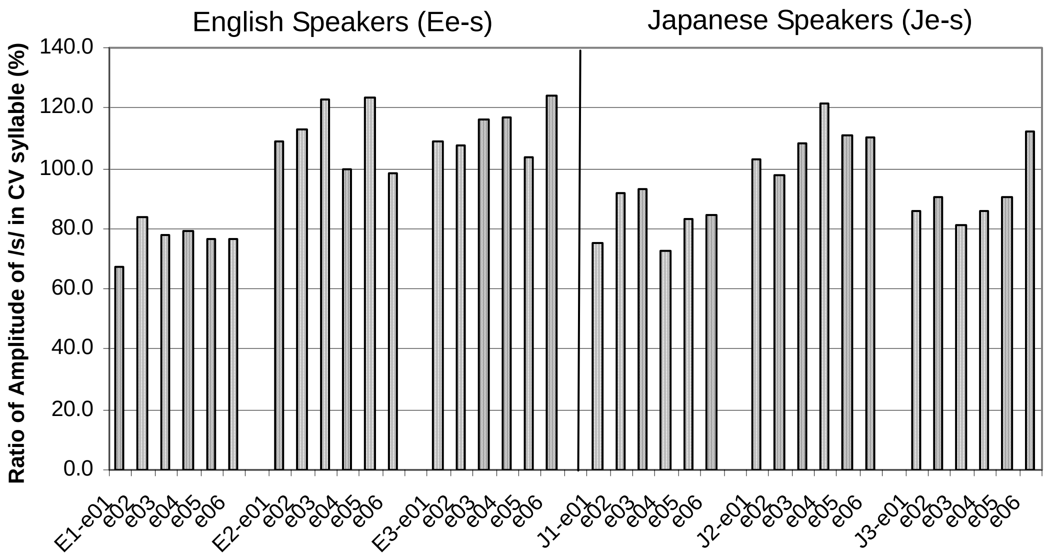
To examine the significance, a statistical test was carried out through analysis of variance (ANOVA). In order to conduct the statistical analysis using ANOVA, all the tokens of /s/ and /θ/ were grouped according to the speakers' native language (i.e. English or Japanese) and the L1/L2 context (i.e. English or Japanese). The mean of acoustic measurements (i.e. ratio of amplitude, proportion of duration, and noise frequency) were then tested for significant differences using ANOVA. The ANOVA produces a value for probability (*p*) which reveals whether mean values are significantly different. The resultant F-ratio and probabilities will be shown. A value of *p* less than 0.05 indicates significantly different means of any acoustical measurement in the comparison groups. When the difference was found not significant, the group was considered homogeneous. In each case the point of interest is whether there is an association between variables. Significance levels will be shown with the marks in Table 6.

Table 6. Significance Levels

P-value (P: probability)	Mark	Significance Level
$p > .10$	ns	Not Significant
$.05 < p < .10$	†	Inclined to be Significant
$p < .05$	††	Significant

5-1. Amplitude

The ratio of noise amplitude of each fricative token to the following vowel was calculated. Figures 10, 11 and 12 display the ratio of amplitude of /s/ in the English word list, the ratio of amplitude of /s/ in the Japanese word list, and the ratio of amplitude of /θ/ in the English word list respectively.



---

[9] The statistical results for the other three cases are as follows; [F(1, 34)=1.81, ns] for (2), [F(1, 34)=0.23, ns] for (3), and [F(1, 34)=1.52, ns] for (4).

Figure 10. Ratio of amplitude of /s/ in English WL

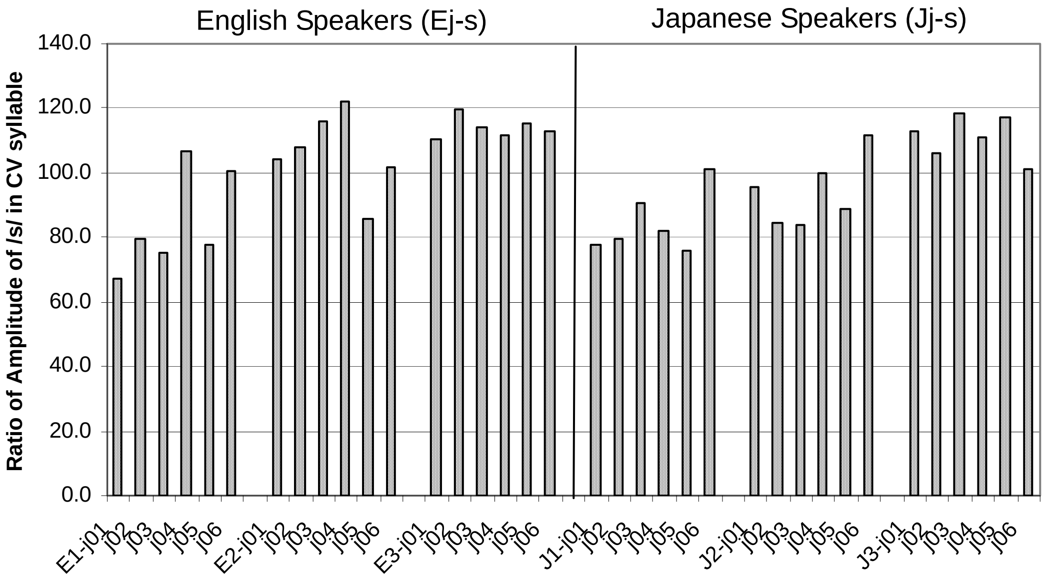


Figure 11. Ratio of amplitude of /s/ in Japanese WL

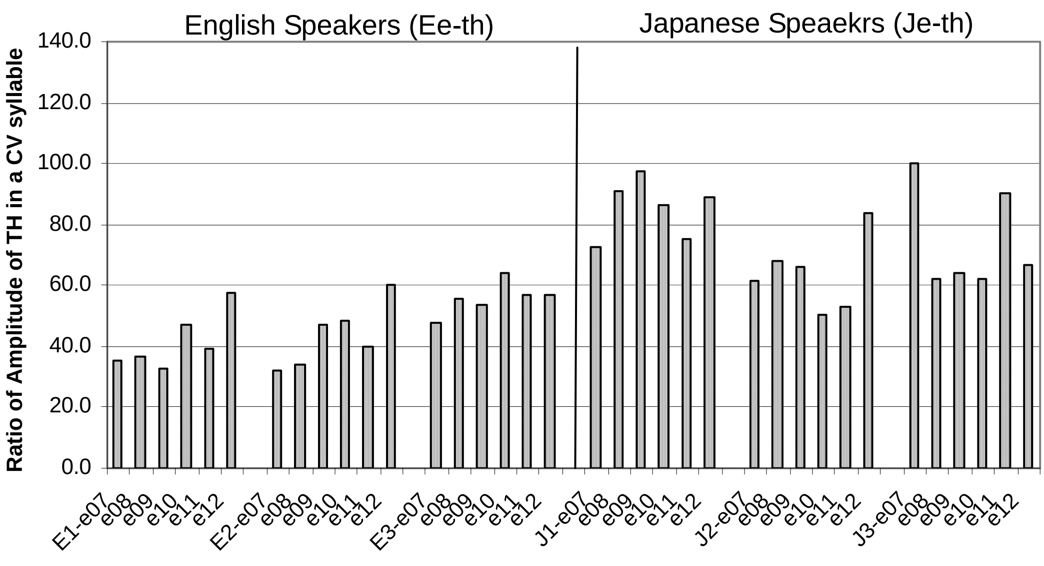


Figure 12. Ratio of amplitude of /θ/ in English WL

Table 7 shows the mean values and S.D. (standard deviation) for the ratio of amplitude of /s/ and /θ/ by the speakers (i.e. English or Japanese) and contexts (L1 or L2), and Figure 13 shows their mean values.

Table 7. Mean values and S.D. for Ratio of Amplitude by speakers and contexts

Speakers	English (E1-3)			Japanese (J1-3)		
Contexts	English (L1)		Japanese(L2)	English (L1)		Japanese(L2)
Phoneme	/s/	/θ/	/s/	/s/	/θ/	/s/
Ref.	Ee-s	Ee-th	Ej-s	Je-s	Je-th	Jj-s
N	18	18	18	18	18	18
Mean	100.2	47.0	101.6	94.4	74.5	96.6
S.D.	18.2	10.1	16.4	13.4	16.0	14.1

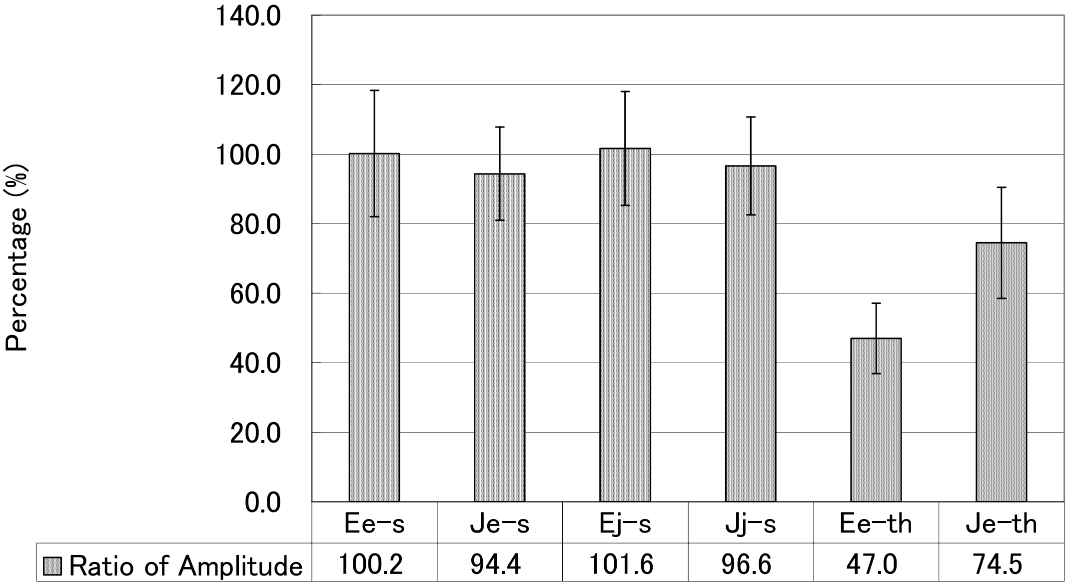


Figure 13. Mean Values and S.D. for Ratio of Amplitude

All these mean values are compared with each other through ANOVA in the following ways:

(1) English /s/ by English (L1) speakers vs. Japanese /s/ by Japanese (L1) speakers, i.e. Ee-s vs. Jj-s
(2) English /s/ by English (L1) speakers vs. English /s/ by Japanese (L2) speakers, i.e. Ee-s vs. Je-s,
(3) Japanese /s/ by English (L2) speakers vs. Japanese /s/ by Japanese (L1) speakers, i.e. Ej-s vs. Jj-s, and
(4) English /θ/ by English (L1) speakers vs. English /θ/ by Japanese (L2) speakers,

i.e. Ee-th vs. Je-th.

The ANOVA reveals that there are no significant differences between any two groups in the cases of (1), (2), and (3): [$F(1, 34)=0.37$, ns] for (1), [$F(1, 34)=1.12$, ns] for (2), and [$F(1, 34)=0.87$, ns] for (3). This suggests there is no evidence of any difference in amplitude between /s/ by the English and /s/ by the Japanese in their L1 and L2 contexts, apart from the fact that both differ from English speakers' /θ/ and to a lesser extent Japanese speakers' /θ/.

The ANOVA reveals that only the comparison in the case of (4) produces a significant result; that is, the average ratio of amplitude of /θ/ produced by the Japanese is significantly higher than that produced by the English [$F(1, 34)=39.81$, $^{\dagger\dagger}p<0.05$]. It may be because some of the /s/ tokens uttered by the Japanese speakers were not the dental realisations [θ] but the alveolar realisations [s].

5-2. Duration

The proportion of duration of each fricative token to the following vowel was calculated. Figures 14, 15 and 16 show the proportion of /s/ duration in English WL, the proportion of /s/ duration in Japanese WL, and the proportion of /θ/ duration in English WL respectively.

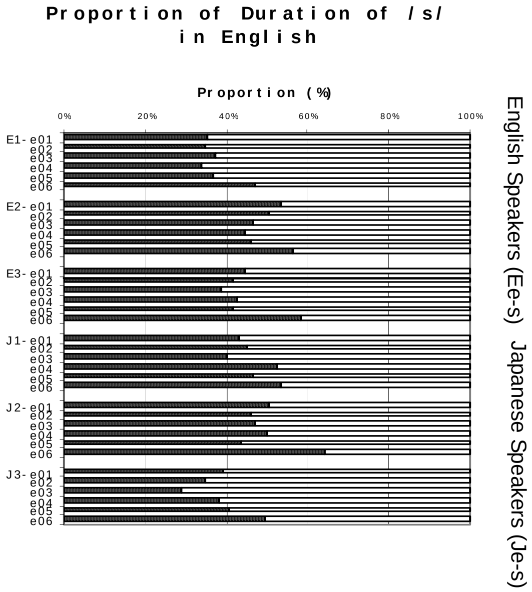


Figure 14. Proportion of /s/ duration in English WL

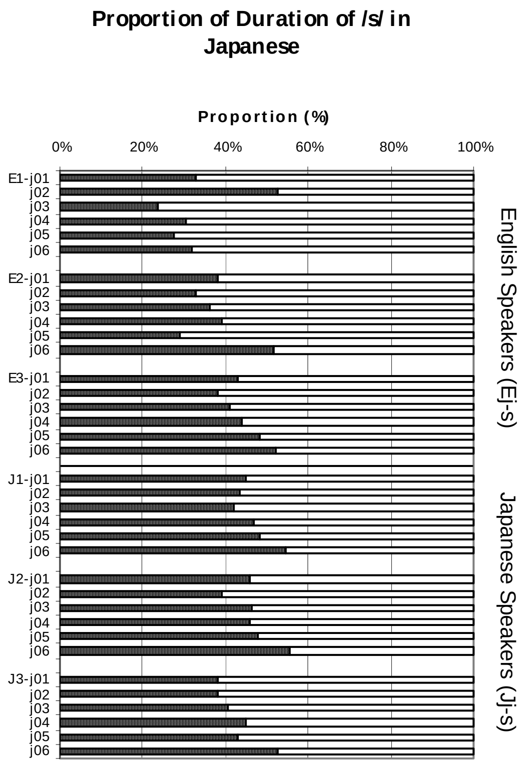


Figure 15. Proportion of /s/ duration in Japanese WL

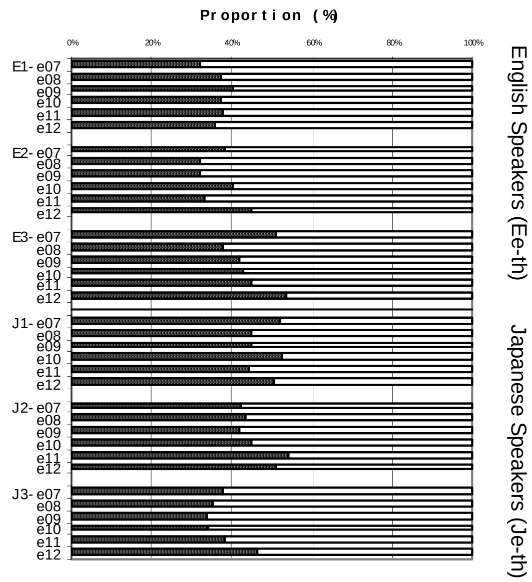


Figure 16. Proportion of /θ/ duration in English WL

Table 8 shows the mean values and S.D. (standard deviation) for the proportion of duration of /s/ and /θ/ by the speakers (i.e. English or Japanese) and contexts (L1 or L2), and Figure 17 shows their mean values.

Table 8. Mean Values and S.D. for Proportion of Duration

Speakers	English (E1-3)			Japanese (J1-3)		
Context	English (L1)		Japanese(L2)	English (L1)		Japanese(L2)
Phoneme	/s/	/θ/	/s/	/s/	/θ/	/s/
Ref.	Ee-s	Ee-th	Ej-s	Je-s	Je-th	Jj-s
N	18	18	18	18	18	18
Mean	43.9	39.6	38.5	45.1	44.1	45.5
S.D.	7.2	6.0	8.6	7.7	6.2	5.0

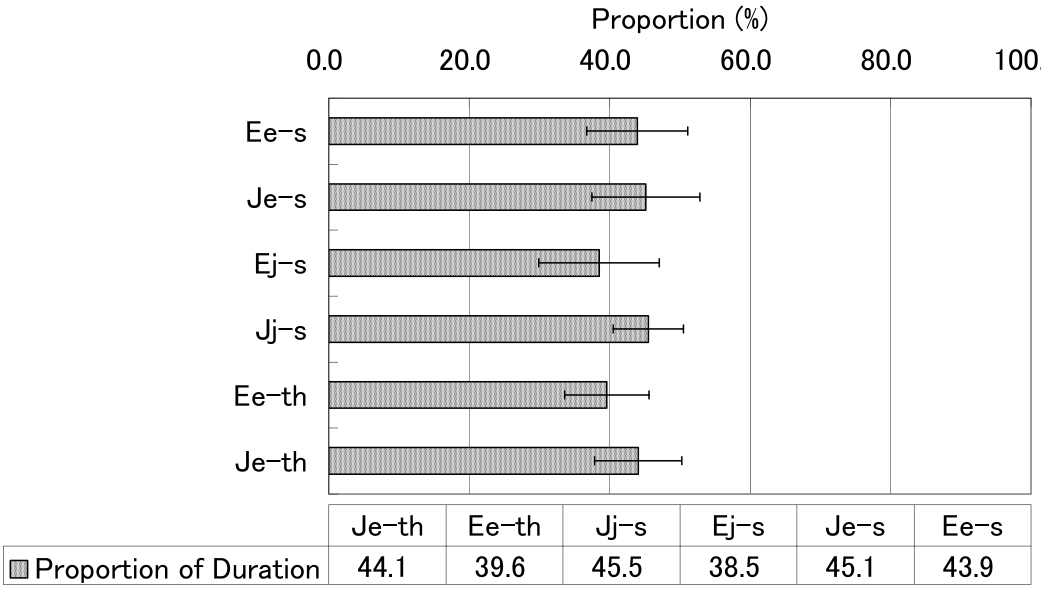


Figure 17. Mean Values and S.D. for Proportion of Duration

All these mean values are compared with each other through ANOVA in the following ways as in the previous section;

(1) English /s/ by English (L1) speakers vs. Japanese /s/ by Japanese (L1) speakers, i.e. Ee-s vs. Jj-s

(2) English /s/ by English (L1) speakers vs. English /s/ by Japanese (L2) speakers, i.e. Ee-s vs. Je-s,

(3) Japanese /s/ by English (L2) speakers vs. Japanese /s/ by Japanese (L1) speakers, i.e. Ej-s vs. Jj-s, and

(4) English /θ/ by English (L1) speakers vs. English /θ/ by Japanese (L2) speakers, i.e. Ee-th vs. Je-th.

Contrary to our expectations, all the mean values of the proportion of duration for /s/ and /θ/ produced by the Japanese speakers are slightly greater than those by English speakers in the all cases. However, according to the ANOVA, the comparison is not always significant; the two groups in the cases of (1) and (2) do not produce any significant results, i.e. [F(1, 34)=0.56, ns] for (1) and [F(1, 34)=0.23, ns] for (2), while the differences between the groups in the cases of (3) and (4) are significant, i.e. [F(1, 34)= 8.54, ††p<0.05] for (3) and [F(1, 34)=4.61, ††p<0.05] for (4). That is, as for /s/, on the one hand, the proportion of /s/ duration of Japanese speakers is significantly greater than that of English speakers only when they are compared in the Japanese context. As to /θ/, on the other hand, the proportion of /θ/ duration by Japanese speakers is significantly greater than that of English speakers; the reason for this may be, as in §5-1, because some of the /θ/ tokens uttered by Japanese speakers were not the dental realisations [θ] but the alveolar realisations [s].

5-3. Frequency

5-3-1. Noise frequency peak

The diagram in Figure 18 was constructed from data of (1) in Table 5, showing the highest and lowest frequency of noise peak of each fricative in each WL uttered by each speaker as well as the calculated average frequency of them. The Figure tells us that all the frequency peaks of /s/ range somewhere around 6,000-10,000Hz, while the range of the frequency peak of /θ/ is much wider from 500 up to 10,000Hz. This reflects the characteristics of the spectrum of /s/ and /θ/, ie. the presence/absence of a concentration of energy in a particular frequency region, which we have seen in §2-3-3-3.

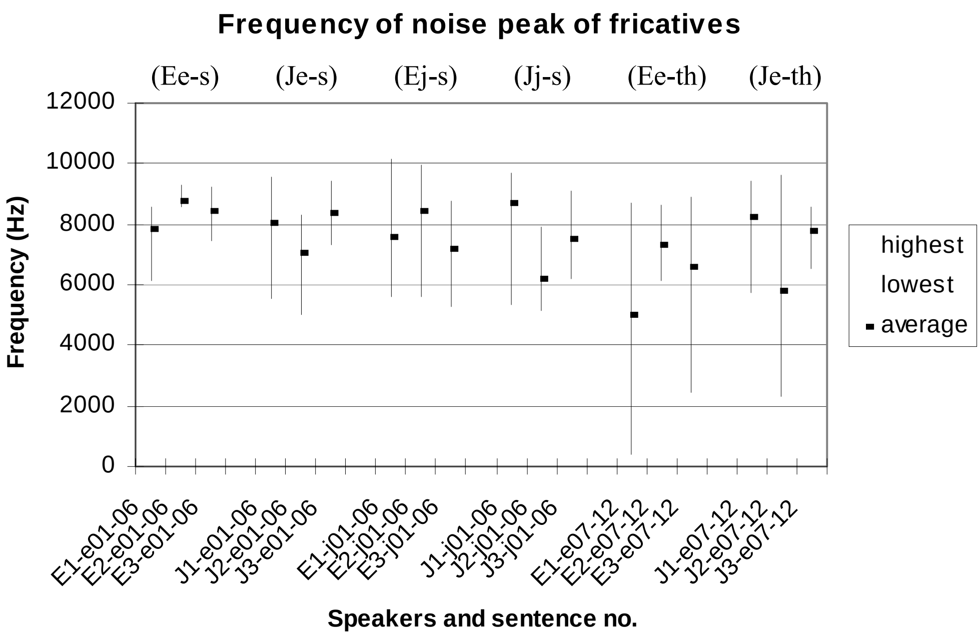


Figure 18. Frequency of noise peak

Table 9 and Figure 19 show the mean values and S.D. (standard deviation) for the frequency of noise peak of /s/ and /θ/ by the speakers (i.e. English or Japanese) and the contexts (L1 or L2).

Table 9. Mean Values and S.D. for Noise Frequency Peak

Speakers	English (E1-3)			Japanese (J1-3)		
Context	English (L1)		Japanese(L2)	English (L1)		Japanese(L2)
Phoneme	/s/	/θ/	/s/	/s/	/θ/	/s/
Ref.	Ee-s	Ee-th	Ej-s	Je-s	Je-th	Jj-s
N	18	18	18	18	18	18
Mean	8351	6313	7742	7815	7271	7474
S.D.	769	2649	1581	1451	1795	1666

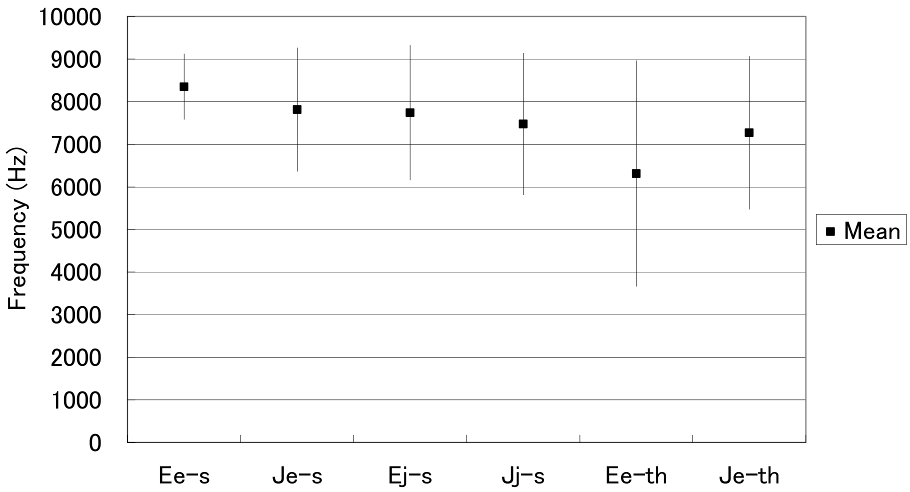


Figure 19. Mean Values and S.D. for Noise Frequency Peak

All these mean values are compared with each other through ANOVA in the following ways;

(1) English /s/ by English (L1) speakers vs. Japanese /s/ by Japanese (L1) speakers, i.e. Ee-s vs. Jj-s

(2) English /s/ by English (L1) speakers vs. English /s/ by Japanese (L2) speakers, i.e. Ee-s vs. Je-s,

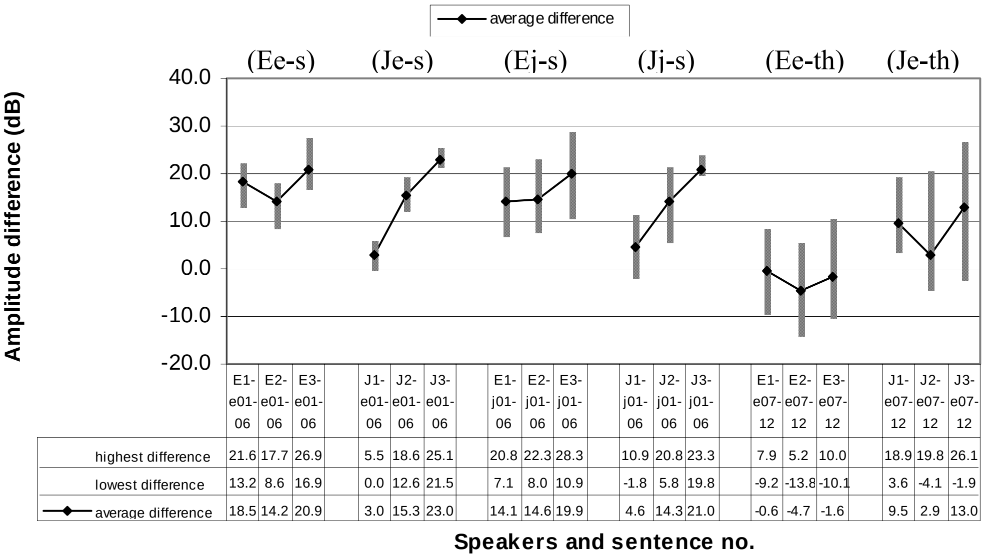
(3) Japanese /s/ by English (L2) speakers vs. Japanese /s/ by Japanese (L1) speakers, i.e. Ej-s vs. Jj-s, and

(4) English /θ/ by English (L1) speakers vs. English /θ/ by Japanese (L2) speakers, i.e. Ee-th vs. Je-th.

As a result, the ANOVA reveals that only the comparison in the first case produces a rather significant result [F(1, 34)=3.88, $^{\dagger}p<0.1$][9]; that is, the mean value for the frequency of noise peak of English /s/ by the English speakers is rather significantly higher than that of English /s/ by the Japanese speakers.

5-3-2. Amplitude difference of the frequency of the region before the following vowel's F5

Figure 20 was constructed from data of (10) in Table 5 which shows the difference in amplitude between the frication noise for the fricative before the following vowel's F5 and the vowel's F5. The vertical line shows the range of difference for each fricative in each WL uttered by each speaker, while the dot indicates the average difference for it.



	E1-e01-06	E2-e01-06	E3-e01-06	J1-e01-06	J2-e01-06	J3-e01-06	E1-j01-06	E2-j01-06	E3-j01-06	J1-j01-06	J2-j01-06	J3-j01-06	E1-e07-12	E2-e07-12	E3-e07-12	J1-e07-12	J2-e07-12	J3-e07-12
highest difference	21.6	17.7	26.9	5.5	18.6	25.1	20.8	22.3	28.3	10.9	20.8	23.3	7.9	5.2	10.0	18.9	19.8	26.1
lowest difference	13.2	8.6	16.9	0.0	12.6	21.5	7.1	8.0	10.9	-1.8	5.8	19.8	-9.2	-13.8	-10.1	3.6	-4.1	-1.9
average difference	18.5	14.2	20.9	3.0	15.3	23.0	14.1	14.6	19.9	4.6	14.3	21.0	-0.6	-4.7	-1.6	9.5	2.9	13.0

Figure 20. Amplitude difference of the frequency region before the following vowel's F5

All the alveolar fricatives uttered by English speakers are more than 10 dB higher in amplitude than the following vowels at the region of the vowels' F5, while all the dental fricatives are lower in amplitude than the vowels, or even if higher, they are no more than 10 dB higher than the following vowels. In the case of the fricatives uttered by Japanese speakers, however, it seems that there is not such a clear tendency for all the three speakers in general. In the case of J2 and J3, the amplitude difference for their alveolar fricatives is comparatively higher than that for their dental fricatives. In the case of J1, however, it is comparatively lower than that for the dental fricatives; this J1's /s/ in both English and Japanese contexts looks more similar to the differences of English speakers' /θ/. For this reason, the alveolar realisations of J1 could possibly be considered to be dentalised, although auditorily they sound like alveolar realisations.

Table 10 and Figure 21 show the mean values and S.D. (standard deviation) for the amplitude difference of the frequency region before the following vowel's F5 by the speakers (i.e. English or Japanese) and the contexts (L1 or L2).

Table 10. Mean Values and S.D. for the Amplitude Difference of the Frequency Region before the Following Vowel's F5

Speakers	English (E1-3)			Japanese (J1-3)		
Context	English (L1)		Japanese(L2)	English (L1)		Japanese(L2)
Phoneme	/s/	/θ/	/s/	/s/	/θ/	/s/
Ref.	Ee-s	Ee-th	Ej-s	Je-s	Je-th	Jj-s
N	18	18	18	18	18	18
Mean	17.8	-2.3	16.2	13.8	8.5	13.3
S.D.	4.1	7.8	6.1	8.4	8.7	7.9

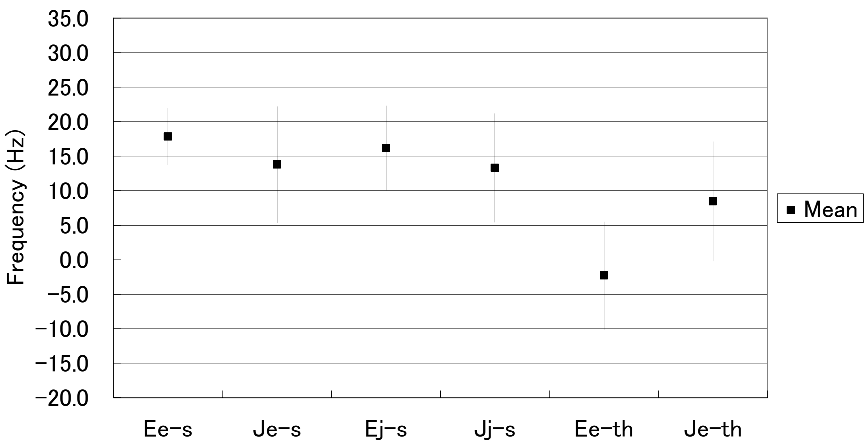


Figure 21. Mean Values and S.D. for the Amplitude Difference of the Frequency Region before the Following Vowel's F5

All these mean values are compared with each other through ANOVA in terms of;

(1) English /s/ by English (L1) speakers vs. Japanese /s/ by Japanese (L1) speakers, i.e. Ee-s vs. Jj-s

(2) English /s/ by English (L1) speakers vs. English /s/ by Japanese (L2) speakers, i.e. Ee-s vs. Je-s,

(3) Japanese /s/ by English (L2) speakers vs. Japanese /s/ by Japanese (L1) speakers, i.e. Ej-s vs. Jj-s, and

(4) English /θ/ by English (L1) speakers vs. English /θ/ by Japanese (L2) speakers, i.e. Ee-th vs. Je-th.

As a result, the ANOVA reveals that the comparison in the cases of (1) and (4) produces a significant result (i.e. [F(1, 34)=4.43, $^{††}p<0.05$] for (1) and [F(1, 34)=14.45, $^{††}p<0.05$] for (4)), the comparison in the case of (2) a rather significant result (i.e. [F(1, 34)=3.16, $^{†}p<0.1$]), and the comparison in the case of (3) a non-significant result (i.e. [F(1, 34)=1.41, ns]); that is, the mean value for the English /s/ by the English speakers is significantly higher than that of Japanese /s/ by Japanese speakers, and is rather significantly higher than that of English /s/ by the Japanese speakers, and the mean value for the English /θ/ by the English speakers is significantly lower than that of English /θ/ by the Japanese speakers, while the mean value for the Japanese /s/ by English speakers and that for the Japanese /s/ by the Japanese speakers are not significantly different.

Given the result assuming that the alveolar realisations of J1 could be considered to be dentalised so that the front cavity could be smaller in space and shorter in length, we could expect that the average frequencies of the lowest spectral peak for J1's alveolar realisations (e01-e06, j01-j06) were higher than those for others; however, they were not always really so. It might possibly be because, as may also be the case for the noise frequency peak that we have already seen above, only one frequency peak cannot tell us anything more important for the nature of fricatives. Instead, we may need to look at overall shape of their spectra.

In fact, if we look at all the spectra, we can find an interesting tendency that can be seen only in the spectra of alveolar fricatives uttered by the English speakers but not in those uttered by the Japanese speakers. Now let us have a look at Figure 22 which shows three spectra of /s/ in the sentence e01 uttered by all three English speakers and Figure 23 which also shows three spectra of /s/ in the sentence j01 (i.e. the corresponding Japanese sentence to the English e01) uttered by all three Japanese speakers.

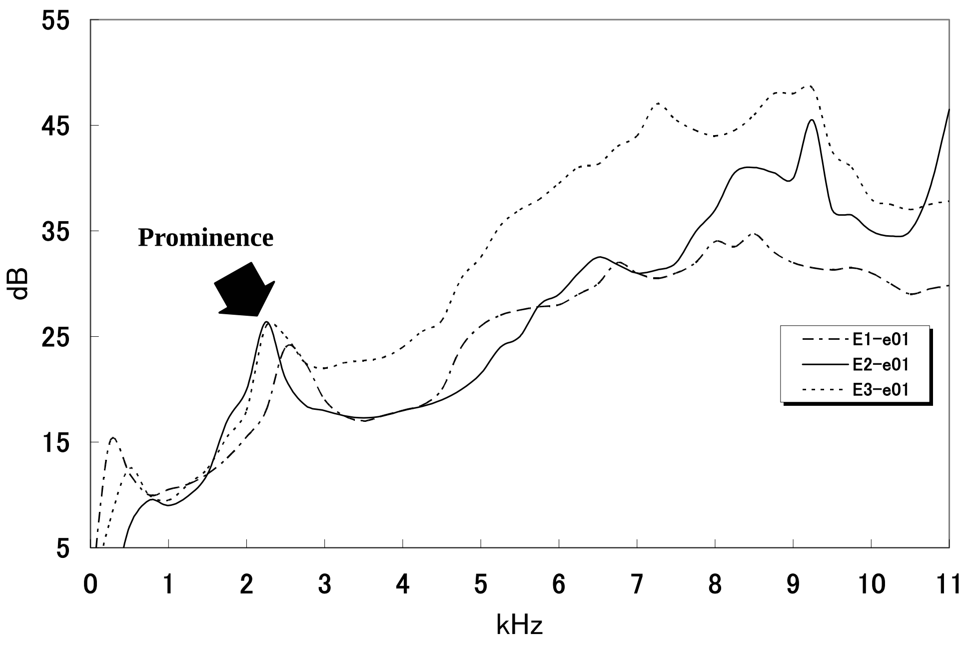


Figure 22. Three spectra of /s/ in the sentence e01 uttered by speakers E1, E2 and E3

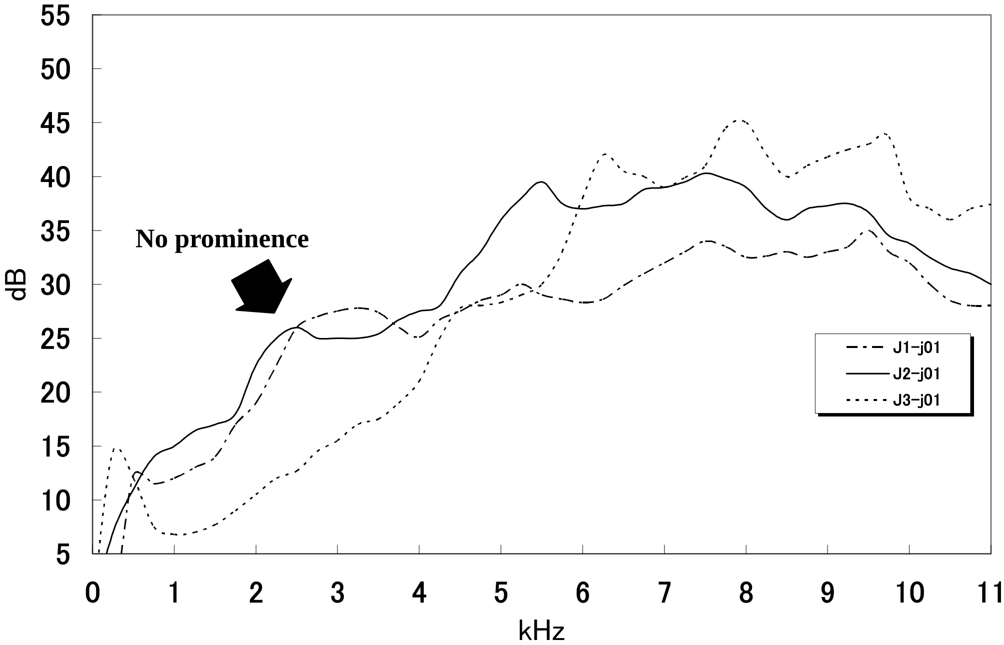


Figure 23. Spectra of /s/ in the sentence j01 uttered by speakers J1, J2 and J3

If we look at these figures carefully, we will find not only that the spectra for the English speakers are more compact at around 8-9kHz than those for the Japanese, but also that the spectra for the English speakers have another prominent energy region at lower frequency, i.e. around 1-3kHz with a significant drop immediately after the prominence, while those for the Japanese do not have this kind of prominence in that frequency region. This tendency also applies to most spectra of the alveolar realisations as shown in Figure 24 and 25 (E1-e06, E2-e06 and E3-e06 vs. J1-j06, J2-j06, and J3-j06).

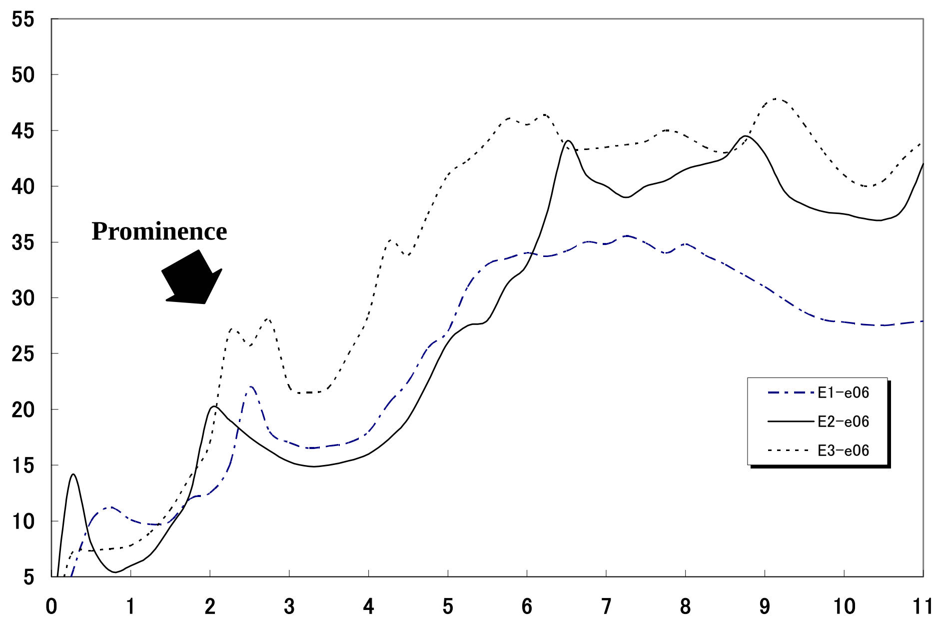


Figure 24. Spectra of /s/ in the sentence e06 uttered by speakers E1, E2 and E3

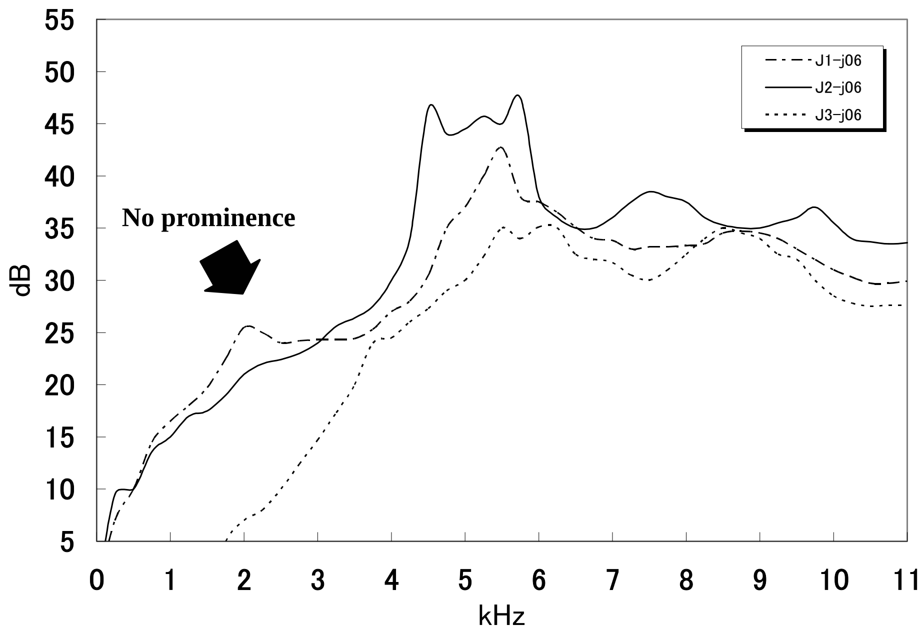


Figure 25. Spectra of /s/ in the sentence j06 uttered by the speakers J1, J2, and J3

6. Conclusions and further studies

As we have seen in the previous sections, we could not obtain significant differences that we had expected in the earlier sections of this paper in terms of intensity, duration, and noise frequency peak. However, we could find one plausible tendency in spectral shape, especially the prominence at the lower frequency region, that could be a crucial acoustic distinction between English /s/ and Japanese /s/. If so, the existence of the prominent energy region at around 1-3kHz in the spectrum of the frication noise of a voiceless alveolar fricative may provide the auditorily stronger impression to the fricatives themselves.

However, these acoustical (i.e. spectral) evidences that could be a potential difference between English /s/ and Japanese /s/ are not enough to confirm that they have something to do with their perception since, as Johnson (2003: 55) said, acoustic analyses give only a rough approximation to the auditory representations that listeners use in identifying speech sounds. Therefore, in the future study, we still need to do a perceptual experiment to investigate the perceptual impression of these sounds, using controlled synthetic speech.

Voiceless alveolar fricatives are often idiosyncratic; therefore it is no wonder if we cannot obtain many general tendencies, as we expected, in such a limited number of tokens. Instead, it should be hoped that the one finding in this paper could shed light on the acoustic description of fricatives which still has considerable room for improvement.

Acknowledgement
I wish to thank my supervisor, Dr. Barry Heselwood, at the Department of Linguistics & Phonetics of the University of Leeds for various comments on earlier drafts of this paper. I am also indebted to an anonymous referee for criticisms and advice. Responsibility for any remaining errors is my own.

References

Cruttenden, A. (2001). *Gimson's Pronunciation of English.* 6th edn. London: Arnold.

Fromkin, V. A. (ed.) (1985). *Phonetic Linguistics: Essays in Honour of Peter Ladefoged.* Orland, FL: Academic Press.

Hayward, K. (2000). *Experimental Phonetics*. London: Longman.

Heinz, J. M. and Stevens, K. N. (1961). On the Properties of Voiceless Fricative Consonants. *Journal of the Acoustical Society of America* **33**. 589-596.

Jacobson, R., Fant, C. G. M. and Halle, M. (1952). *Preliminaries to Speech Analysis*. Acoustics Laboratory, Massachusetts Institute of Technology.

Johnson, K. (2003). *Acoustic & Auditory Phonetics*. 2nd edn. Cambridge, MA and Oxford: Blackwell.

Johnston, P. (1997). Regional Variation. In: Jones, C. (ed.)*The Edinburgh History of the Scots Language.* Edinburgh: Edinburgh University Press. pp.433-513.

Jones, C. (Ed.) (1997). The Edinburgh History of the Scots Language. Edinburgh: Edinburgh University Press.

Kent, R. D. and Read, C. (1992). *The Acoustic Analysis of Speech*. London: Whurr Publishers.

Ladefoged, P. (1993). *A Course in Phonetics*. 3rd edn. New York: Harcourt Brace

College Publishers.
Lambacher, S., Martens, W., Nelson, B. & Berman, J. (2001) Identification of English Voiceless Fricatives by Japanese Listeners: The Influence of Vowel Context on Sensitivity and Response Bias. *Acoustic Science & Technology*. **22**.
Laver, J. (1994). Principles of Phonetics. Cambridge: Cambridge University Press.
Macafee, C. (1983). *Traditional Dialect in the Modern World: A Glasgow Case Study*. Frankfurt: Peter Lang.
Nartey, J. N. A. (1982). On Fricative Phones and Phonemes: Measuring the Phonetic Differences within and between Languages. *UCLA Working Papers in Phonetics* No. 55.
Ono, S. (1986). *Eigo Onseigaku Gairon* [An Outline of English Phonetics]. Tokyo: Liber Press.
Saito, H. and Takebayashi, S. (1998). *Eigo Onseigaku Nyuumon* [An Introduction to English Phonetics]. Tokyo: Kenkyusha.
Stevens, K. N. (1985). Evidence for the Role of Acoustic Boundaries in the Perception of Speech Sounds. In: Fromkin, V. A. (ed.) *Phonetic Linguistics: Essays in Honour of Peter Ladefoged.* Orland, FL: Academic Press. 243-255.
Stuart-Smith, J. (2004). Empirical Evidence for Gendered Speech Production: /S/ in Glaswegian. Paper presented at *Laboratory Phonology* ***9***, 24-26 June 2004, University of Illinois.
Takebayashi, S. (1996). *Eigo Onseigaku* [English Phonetics]. Tokyo: Kenkyusha.
Tanaka, S. (1996). *Jissen Shinri Deeta Kaiseki* [Practical Analysis of Psychological Data]. Tokyo: Shinyosha.
Trask, R. L. (1996). *A Dictionary of Phonetics and Phonology*. London: Routledge.
Wells, J. C. (2000). *Longman Pronunciation Dictionary*. 2nd edn. Harlow: Longman.
You, H. Y. (1979). *An Acoustical and Perceptual Study of English Fricatives*. Unpublished MA dissertation, University of Edmonton, Edmonton, Canada.

Miho Kamata
Department of Linguistics & Phonetics
University of Leeds
LS2 9JT

lnp2mk@leeds.ac.uk

Appendix 1. Measured values of acoustic components (1)-(10) for /s/ and /θ/

(1): Frequency peak of the average spectrum of fricatives (Hz)
(2): Amplitude of (1) (dB)
(3): Frequency peak of the average spectrum of the vowels following fricatives (Hz)
(4): Amplitude of (3) (dB)
(5): Duration of fricatives (ms)
(6): Duration of the following vowels (ms)
(7): F5 of the following vowels (first 50 ms) (Hz)
(8): Average Amplitude of the frequency region of (7) of fricatives (last 50 ms) (dB)
(9): Average Amplitude of (7) (dB)
(10): difference between (8) and (9) (dB)

Speaker	No.	phoneme	realisation	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
E1	e01	s	s	8232	34.7	616	51.6	163	296	7891	31.2	18	13.2
E1	e02	s	s	8507	37.8	1511	45.2	140	264	9066	35.8	16.7	19.1
E1	e03	s	s	8563	35	782	45.1	158	264	7947	31.6	13.2	18.4
E1	e04	s	s	6156	33.4	784	42.3	149	291	8227	32.4	10.8	21.6
E1	e05	s	s	8171	34	784	44.4	156	268	7891	32.7	12.7	20
E1	e06	s	s	7275	35.6	448	46.5	213	238	7947	31.9	13.3	18.6
E1	e07	θ	θ	392	19.2	616	54.5	140	293	7499	11.3	20.5	-9.2
E1	e08	θ	θ	1024	17.9	1511	48.5	179	301	7107	11.7	19.1	-7.4
E1	e09	θ	θ	3190	16.2	1511	50	181	266	5820	10.3	18.2	-7.9
E1	e10	θ	θ	8283	22.2	728	46.9	175	295	7387	20.6	12.7	7.9
E1	e11	θ	θ	8674	18.7	672	47.3	168	279	7891	16.4	11	5.4
E1	e12	θ	θ	8563	26.3	448	45.7	155	279	7779	24.4	16.8	7.6
E1	j01	s	s	10130	32.8	616	48.6	144	297	7499	26.9	17.5	9.4
E1	j02	s	s	9290	34.4	1567	43.1	340	305	8842	31.6	11.9	19.7
E1	j03	s	s	8786	34.9	672	46.4	104	338	8283	34.8	14	20.8
E1	j04	s	s	5596	45.2	839	42.4	131	301	8339	29.2	22.1	7.1
E1	j05	s	s	5596	34.6	616	44.5	118	307	10018	25.7	11.4	14.3
E1	j06	s	s	6044	40.3	336	40.1	135	288	9794	29.2	15.9	13.3
E2	e01	s	s	9290	44	2742	40.4	169	146	7611	32.7	20.4	12.3
E2	e02	s	s	8563	51.7	1567	45.9	187	183	7052	38.5	24.5	14
E2	e03	s	s	8674	48.7	686	39.7	146	169	7499	41.7	24	17.7
E2	e04	s	s	8786	43.6	839	43.8	169	210	7499	38.2	29.6	8.6
E2	e05	s	s	8563	49.2	1343	39.8	179	210	7723	41.2	25.6	15.6
E2	e06	s	s	8730	43.7	2462	44.4	245	189	7499	38.9	22.1	16.8
E2	e07	θ	θ	6156	14.7	560	46.1	111	181	8003	11.1	20.8	-9.7
E2	e08	θ	θ	6884	15	895	43.9	90	192	7891	15.2	16.8	-1.6
E2	e09	θ	θ	8619	22	839	46.4	97	205	7779	22.6	17.4	5.2
E2	e10	θ	θ	6660	23.4	839	48.1	142	212	7611	22.3	31.9	-9.6
E2	e11	θ	θ	7835	17.3	839	43.2	130	260	7499	13.9	27.7	-13.8
E2	e12	θ	θj	7835	23.1	2910	38.5	162	200	7275	22.3	21.1	1.2
E2	j01	s	s	7835	39.2	2742	37.7	99	159	7443	30.7	19.2	11.5

E2	j02	s	s	8842	43	672	39.8	101	205	7163	30.9	21.1	9.8
E2	j03	s	s	9962	48	784	41.4	154	273	9010	39.6	17.3	22.3
E2	j04	s	s	8954	41.4	728	33.9	139	217	7667	37.8	17.8	20
E2	j05	s	s	8730	28.3	448	33	126	306	7443	28	20	8
E2	j06	s	s	6268	39	336	38.4	182	169	7331	37.9	22.1	15.8
E3	e01	s	s	9234	48.5	560	44.6	185	228	7443	44.3	27.1	17.2
E3	e02	s	s	9066	46.2	1175	43	174	245	7611	37.7	17.9	19.8
E3	e03	s	s	8059	48.2	895	41.5	153	243	7387	45.8	24.7	21.1
E3	e04	s	s	7443	49.3	504	42.1	159	215	7387	50.3	23.4	26.9
E3	e05	s	s	7835	45.9	616	44.4	168	236	6716	43.5	20.3	23.2
E3	e06	s	sj	9178	47.8	336	38.6	221	157	9570	42.3	25.4	16.9
E3	e07	θ	θ	6044	19.4	616	40.6	226	215	7443	13.5	23.6	-10.1
E3	e08	θ	θ	5820	24.5	395	44.2	166	273	7555	18.6	28	-9.4
E3	e09	θ	θ	2462	20.3	839	37.9	183	257	7499	14.7	24.2	-9.5
E3	e10	θ	θ	8059	26.2	504	40.7	168	226	7723	24.5	14.5	10
E3	e11	θ	θ	8227	23	616	40.6	218	269	7611	16.5	13.4	3.1
E3	e12	θ	θ	8898	22.3	336	39.1	228	197	7052	18.5	12	6.5
E3	j01	s	s	8786	41.6	448	37.6	154	204	10353	30.7	18.7	12
E3	j02	s	s	8395	44.7	784	37.3	143	233	8395	48.5	20.2	28.3
E3	j03	s	s	8115	42.7	1007	37.5	192	275	8059	43.3	21.5	21.8
E3	j04	s	s	7219	41.3	448	36.9	192	245	7835	36.6	14.6	22
E3	j05	s	s	5261	43.4	504	37.6	227	245	7835	38.3	14.2	24.1
E3	j06	s	s	5540	40.4	336	35.8	220	202	9794	24.3	13.4	10.9
J1	e01	s	s	9234	35.7	504	47.5	136	179	8730	21.8	21.8	0
J1	e02	s	s	9010	37.2	384	40.6	157	192	6828	34.6	32.4	2.2
J1	e03	s	s	9122	36.6	392	39.4	159	237	7387	36.1	30.6	5.5
J1	e04	s	s	5652	35.8	560	49.1	207	187	7331	32.7	28.1	4.6
J1	e05	s	s	5540	37	556	44.3	179	207	7723	31.9	29.4	2.5
J1	e06	s	s	9570	38	336	45	192	167	7107	31.5	28.1	3.4
J1	e07	θ	s	9234	30.3	448	41.6	182	169	7891	29.4	25.8	3.6
J1	e08	θ	s	9402	35.4	448	38.9	172	212	7667	32.6	24.1	8.5
J1	e09	θ	s	8619	35.2	384	36.1	162	200	7611	30.6	21.7	8.9
J1	e10	θ	s	8059	41.8	560	48.4	207	187	7555	37.8	18.9	18.9
J1	e11	θ	s	5764	33.7	616	44.6	154	192	7443	28.4	23	5.4
J1	e12	θ	s	8227	36.7	392	41.2	182	177	8730	31.4	19.6	11.8
J1	j01	s	s	9514	35.1	504	45	114	139	7387	35	30.3	4.7
J1	j02	s	s	9682	34.1	504	42.7	114	149	7163	34.8	36.6	-1.8
J1	j03	s	s	9514	38.1	560	41.9	116	159	7499	35.7	33.1	2.6
J1	j04	s	s	8954	37.5	560	45.6	136	154	7611	36.9	31.1	5.8
J1	j05	s	s	9290	34.5	504	45.4	129	139	7555	34.9	29.3	5.6

J1	j06	s	s	5317	41.9	392	41.4	162	134	7387	33.4	22.5	10.9
J2	e01	s	s	8283	43.5	2970	42.3	152	150	9514	36.9	18.3	18.6
J2	e02	s	s	7997	39.4	685	40.3	149	174	8171	30.9	18.3	12.6
J2	e03	s	s	7997	40.4	914	37.4	158	179	8115	32.5	15.1	17.4
J2	e04	s	s	5027	43.1	800	35.4	145	145	9682	22.1	8.3	13.8
J2	e05	s	s	7883	40	800	36.1	131	170	8283	33.1	19.3	13.8
J2	e06	s	s	5198	39.6	343	36	186	104	8003	25.5	9.7	15.8
J2	e07	θ	θ	2285	25.7	457	41.6	118	161	8451	24.9	29	-4.1
J2	e08	θ	θ	9597	25.9	685	38	129	170	8003	22.2	26.3	-4.1
J2	e09	θ	θ	5712	25.6	914	38.9	127	178	8059	25.1	21.5	3.6
J2	e10	θ	θ	5655	22.2	800	44.2	127	155	8003	23.3	21.9	1.4
J2	e11	θ	θ	5655	21.1	815	39.7	120	101	7891	18.1	17.6	0.5
J2	e12	θ	θ	5876	32.7	336	39	163	155	8059	34.3	14.5	19.8
J2	j01	s	s	7723	39.7	560	41.5	116	136	9010	32	16.4	15.6
J2	j02	s	s	7891	39.7	784	47.1	116	179	7387	34.5	14.3	20.2
J2	j03	s	s	5540	41	784	48.8	164	189	10018	21.5	14.6	6.9
J2	j04	s	s	5149	41.7	784	41.7	119	139	10074	31.2	10.4	20.8
J2	j05	s	s	5429	42	560	47.2	144	157	8003	22.1	16.3	5.8
J2	j06	s	s	5485	46.8	448	42	146	116	9850	35.9	19.7	16.2
J3	e01	s	s	8569	44	685	51.2	125	195	8674	45.3	20.2	25.1
J3	e02	s	s	9026	43.5	743	48	118	221	8674	44	19.9	24.1
J3	e03	s	s	7940	38.3	685	47	95	235	7947	36.6	15.1	21.5
J3	e04	s	s	9426	38	743	44.2	130	210	8563	36.9	14.5	22.4
J3	e05	s	s	7883	42.8	743	47.4	131	190	9290	40.2	17.5	22.7
J3	e06	s	s	7312	40	400	35.6	136	140	8674	34.6	12.3	22.3
J3	e07	θ	s	8569	48.9	685	48.7	118	194	8563	47.8	21.7	26.1
J3	e08	θ	θ	6512	26.9	743	43.4	127	233	8339	17.9	19.8	-1.9
J3	e09	θ	s	8226	28.4	628	44.3	115	224	8395	26.2	16.7	9.5
J3	e10	θ	s	7826	27.1	685	43.5	126	240	8451	26.8	13.8	13
J3	e11	θ	s	7769	37.3	743	41.3	129	206	7947	33.9	12.1	21.8
J3	e12	θ	θj	7883	26.4	400	39.7	156	179	9066	20.4	10.8	9.6
J3	j01	s	s	8059	44.8	616	39.6	98	159	8954	41.2	17.9	23.3
J3	j02	s	s	6492	42.6	560	40.1	114	184	8730	40.1	19.3	20.8
J3	j03	s	s	9066	47.5	772	40.2	124	182	8395	41.7	21.9	19.8
J3	j04	s	s	6268	40.5	616	36.4	126	154	9682	31.4	10.8	20.6
J3	j05	s	s	8954	42.2	448	36	114	152	8227	33.7	12.4	21.3
J3	j06	s	s	6212	35.4	336	35.1	149	134	7387	28.4	8.2	20.2