



TEACHING ENGLISH PRONUNCIATION TO JAPANESE UNIVERSITY STUDENTS: THE VOICELESS FRICATIVE /S/ SOUND

Hisako Murakawa

International Budo University
Katsuura-shi, Chiba 299-52 Japan

ABSTRACT

It is generally believed that an adult has more difficulty mastering the pronunciation of a second language than does a child. Various studies, from psychological and methodological viewpoints, have found evidence that the child acquires the phonetics of speech much faster than does the adult. Various writers have discussed techniques and methods of teaching classroom. However, in Japan, very few researchers have attempted to verify scientifically whether or not the application of phonetic information in teaching English pronunciation will enhance adult performance.

The purpose of this study is to determine whether or not the phonetic method, used with laser discs which show the articulation of English pronunciation, produces any significant differences from using the audio-lingual method only.

The target English voiceless fricative /s/ sound is chosen. Although the /s/ sound exists in the Japanese language, the English fricative /s/ requires a much stronger breath stream than the Japanese /s/. It is, therefore, one of the most difficult English sounds for Japanese to produce. Learning the correct way of breathing and articulation will facilitate the production of this significant /s/ sound.

Subjects for the experimental group were from the junior class of International Budo University majoring in *budo* (martial ways). The control group are juniors from IBU majoring in physical education. Both groups are in the highest class of English proficiency of the four instructional levels.

An identical measurement test is given before and after the training session. A sound analyzer, one of the devices of PROTS (Pronunciation Training System) is used. All spoken data from both Japanese groups are compared with those of native speakers of American English. The results will be observed.

I. INTRODUCTION

For several years the "Audio-lingual Methods," a teaching method introduced in Japan by Palmer in 1922, was used to improve listening and speaking ability. This method is based on behaviorism [5]. External forces select stimuli and reinforce responses until the desired behavior is conditioned to occur. This method is used in the learning of pronunciation by imitation and repetition of the materials. However, many Japanese adults are still unable to produce adequate English pronunciation with ease. Clearly new methods are needed and more research must be done in methods to improve the teaching and learning of English pronunciation.

This study evaluates the phonetic method of teaching pronunciation. This method uses laser discs which show the differences between Japanese articulation of English and the articulation of a native speaker, phonetic alphabets (the International Phonetic Alphabets), and careful articulatory explanations, using charts and diagrams of the organs of speech

appearing on the screen, are used. The purpose of this study is to determine whether or not the phonetic method produces any significant improvement in pronunciation over the audio-lingual method.

The English voiceless fricative /s/ sound was chosen as the target sound for testing the efficacy of the phonetic method because this sound is not produced clearly and sharply by Japanese speakers. Although the /s/ sound exists in the Japanese language, the English fricative /s/ is much stronger than the Japanese /s/. In other words, the Japanese /s/, alveolar fricative, is articulated with less frication than the English /s/. It is, therefore, one of the most difficult English sounds for the Japanese speakers to produce.

There may be two reasons for this difficulty, articulatory differences between the English and the Japanese /s/, and the differences in the breath energy used to produce the sound. In English, the side of the tongue makes contact with the inner surfaces of the upper back teeth. The blade of the tongue is in contact with the alveolar ridge except for a small V-shaped groove along the midline of the tongue. As the breath is forced through this narrow groove, under great pressure, it strikes the upper edges of the lower front teeth, and produces the sound of friction [1].

In Japanese, the apex of the tongue makes contact with a point between the alveolar ridge and the palato-alveolar area of the roof of the mouth. In articulating the Japanese /s/, the position of the lower jaw of the Japanese speaker is farther back than the position of the jaw of the English speaker when he articulates the English /s/. It may assist the Japanese speaker in producing a clear /s/ if he shifts his lower jaw forward slightly to bring his lower teeth out even with his upper teeth because it is produced farther front in the mouth. If the Japanese speaker follows the above method he will be able to prolong the hiss of the English /s/ very easily.

The duration of the frication noise is another important factor in producing a clear and sharp /s/ sound. Since the fricative /s/ is produced with a narrow constriction in the oral cavity, the duration of the frication noise may influence the perception of fricatives. You (1979) showed that in English the duration of frication noise varies with place of articulation; the frication of alveolo-palatal fricatives averaged 176 ms, and that of alveolars, 155ms [6]. Hughes and Halle (1956) obtained high recognition scores for /s/ using a frication duration of only 50 ms [2]. This study suggests that listeners do not require the frication noise in its entirety for fricative identification. Obviously Hughes and Halle's results suggested that for some voiceless fricatives, duration is not an important perceptual cue for identification only when native English speakers produce and perceive the /s/. In that case a frication duration of only 50 ms is long enough to identify the fricative sounds because fricative sounds are created with the strong breath energy.

The purpose of this study is to observe just how much frication noise is produced by the models (native speakers of American English) and the subjects before and after training the Japanese university students in the American English

pronunciation, and to find what the Japanese students should do in order to produce the adequate American English /s/ sound.

II. PROCEDURES

Measurements were made to compare the experimental group's performance, in terms of oral production, with that of the control group. Both groups' oral production was, then, compared with a criterion group of native speakers of English.

The hypotheses were: The student segment of the experimental group (Group A) was expected to show greater improvement in performance, in terms of production of the English voiceless fricative /s/ sound by the use of the phonetic method, than was the control group (Group B) using only the audio-lingual method. The closer the mean duration of the /s/ sound, the better the quality of intelligibility.

Fourty of the fifty-seven juniors (47 male and 10 female) of the Department of Budo at International Budo University were able to participate in the experimental group. Ten were absent on the day pronunciation was recorded.

Sixty-two juniors (50 male, 15 female students) of the Department of Physical Education at the same university were chosen for the control group. Fifty-five of the sixty-two students were recorded. The rest were absent on the day pronunciation was recorded. The students of both groups were in the highest class of the four English instructional levels of each Department.

Teaching materials for the students in Group A were designed to obtain information relevant to the hypotheses of this study. The devices of PROTS (Pronunciation Training System) were used[4]. PROTS consists of five volumes of textbooks, five sheets of laser discs, and an analyzer. Volume 3, providing detailed information about the English /s/ sound, was used for Group A. The target words and sentences, employing the /s/ sound, were chosen from their volume and given to both groups in the pre and post tests.

The test was designed to monitor the learner's articulation of the /s/ sound. This test covered the allophonic /s/ sound thoroughly, and the sentences employing the /s/ sound provided the opportunity for more variation in oral production errors than is provided by speaking merely one word.

The tests for both groups were given in different class periods of the same week and were recorded in each booth (in the language laboratory).

The subjects were not told that they were participants in an experiment. A spoken announcement informed them that the aim of this diagnostic test was to analyze the errors they made, not to grade them.

Retests, identical to the diagnostic tests, were given to all subjects at the end of training.

Group A and Group B met ninety minutes a day once a week. Only fifteen minutes out of ninety minutes of regular conversation class periods were used for training. The author trained both groups.

In the first class for teaching the /s/ sound to Group A, detailed articulatory explanations were made and shown on the screen using the laser discs. In addition to the method for articulate the /s/ sound, the instruction emphasised the importance of using a strong breath stream and prolonging the /s/ sound when articulating.

First, the teacher asked the students to imitate biting a finger nail, resulting in the student's upper teeth touching their lower teeth. This facilitates the production of the /s/ sound. The teacher moved around the class to make sure their articulation was adequate, and then asked them to repeat words and sentences after the model.

For Group B, the teacher asked students to repeat after the

model recorded from the laser discs into the cassette tape. When the student could not pronounce the correct sound, the teacher asked that it be repeated after the tape a few times more, without giving any explanation of articulators.

The words and sentences were recorded. The syllables in them were consonant-vowel (CV) utterances consisting of the voiceless fricative /s/ followed by one of the vowels [i, I, ε, æ, ^], and a diphthong [aI], vowel-consonant (CV) utterances consisting of the /s/ preceded by one of the vowels [I, ε], and clusters in which the /s/ follows or precedes one of the voiceless stop consonants [s,p,t,k,f]. Each utterances were printed on B4 size paper in voiceprint and wave form. These words and sentences were produced by one female native speaker of American English, a model in the laser disc, and two female and two male native speakers of American English. The naiter speakers read all words and sentences at a normal speaking rate. One female's utterances were recorded from the laser disc onto cassette tape with a TC-1450 tape recorder as well as other American English speakers. The student's utterances were recorded on cassette tape in a language laboratory with a head set (WE 5920) of a booth recorder (WE6500). Stimuli were transferred to a PROTS analyzer for editing.

Measurements were taken directly from a displayed screen by moving two cursors to locate the frication. First, the onset of the fricative /s/ was identified. Second, the offset of the /s/ was identified. For /s/ produced by the Japanese students, this point was sometimes hard to locate.

In measuring the spoken data, noise was unavoidable since the fricative is characterized by a very weak onset and offset of energy as shown in Figure 1 and 2 and in Figure 3 and 4 which show the different energy levels and also all of the students pronounced the target words and sentences at the same time in the language laboratory. Therefore, it was often difficult to identify whether the subject's frication or some noise frication was being registered especially around the low frequency levels. In this case, a high frequency energy level was required in order to set the cursor on the noise-onset. Because of the noise, the wave forms used to compare the students' and model's wave forms were abandoned as a criteria.

FIGURE 1. The weak frication noise of the /s/ sound produced by the Japanese speaker

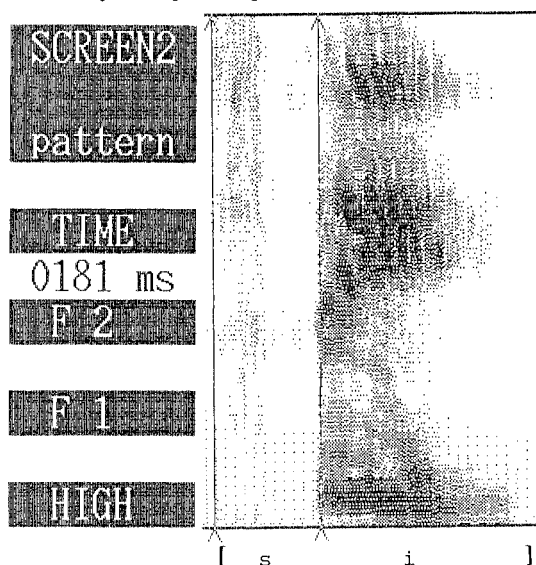


FIGURE 2 The strong frication noise of the /s/ sound produced by the American English speaker

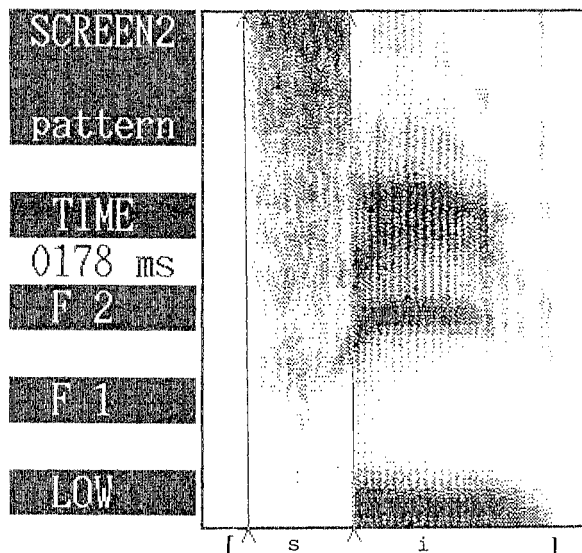
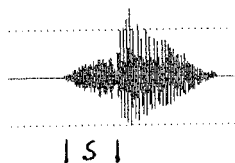


FIGURE 3 The weak breath energy of the /s/ sound produced by the Japanese speaker



FIGURE 4 The strong breathenergy of the /s/ sound produced by the American English speaker



The words and sentences used for this experiment are as follows:

(single word)

/s/ in the initial position: see, sandwich, sixteen

/s/ in the middle position: assignment

/s/ in the final position: pass, face, bus, base, miss

/s/ cluster in the initial position: specia, standard

/s/ cluster in the middle position: sixteen, telescope, textbook

/s/ cluster in the final position: ask, best

(words in sentences)

/s/ in the initial position: see, sister, successful

/s/ in the middle position: businessman

/s/ in the final position: nice, nurse

/s/ cluster in the middle position: successful, successful, sister

It was nice to see you.

My sister is a nurse.

He is a successful businessman.

III ANALYSIS OF RESULTS

Table I shows the duration of the full frication noise of the test stimuli uttered by female and male American English speakers. Speaker A's duration of the [s] in the initial and final positions is shorter than other four speakers'. On the other hand, Speaker A's duration of the [s] cluster in the middle position is longer than other speakers.

TABLE I Duration (in ms) of the full frication of the test stimuli (American English speakers)

Speaker	A(F)	B(F)	C(F)	D(M)	E(M)	Av.
	(Las.D)					
(single word)						
init.	<u>see</u>	181	189	180	185	178 183
	<u>sandwich</u>	145	225	175	196	335 215
	<u>sixteen</u>	132	191	177	188	187 175
mid.	<u>assignment</u>	170	190	175	176	118 166
final	<u>pass</u>	96	232	198	195	187 182
	<u>face</u>	115	315	213	220	185 210
	<u>bus</u>	93	240	189	350	426 260
	<u>base</u>	110	230	216	315	186 211
	<u>miss</u>	132	253	205	330	239 232
/s/ clus.						
init.	<u>special</u>	129	175	183	191	173 170
	<u>standard</u>	151	191	176	185	126 166
mid.	<u>sixteen</u>	96	188	173	178	165 160
	<u>telescope</u>	151	110	98	88	85 106
	<u>textbook</u>	104	172	183	165	178 160
final	<u>ask</u>	50	182	175	193	181 156
	<u>best</u>	154	183	119	189	165 162
(words in sentence)						
init.	<u>see</u>	151	161	152	174	154 158
	<u>sister</u>	129	188	135	163	155 154
	<u>successful</u>	38	173	137	181	110 128
mid.	<u>businessman</u>	118	116	118	116	115 117
final	<u>nice</u>	79	125	115	119	112 110
	<u>nurse</u>	183	215	186	208	203 199
/s/ clus.						
	<u>successful</u>	96	112	103	108	107 105
	<u>successful</u>	110	108	96	102	85 100
	<u>sister</u>	90	102	85	90	85 90

Although Speaker A's duration often indicates less than 100 ms as seen in the words, "pass," "bus," "successful," "nice," and "sister," her noise frequency level is very high. Therefore, her [s] sound is always perceived as a clear and sharp sound. Other four speakers' duration is consistent, except for the words, "telescope," and the second middle [s] sound in "successful" and "sister." The mean duration of the [s] in the final position is longer than that of the initial position. The mean duration of the [s] in the cluster is shorter than that of the initial position. The mean duration of the /s/ uttered in the sentences is the shortest duration of all.

Jongman (1989) provides duration data on the [s] sound in the initial position: 185 ms in [sa], 193 ms in [si], 187 ms in [su] [3]. Although not all the durations of five speakers reach the duration given by Jongman, the duration of the [s] in the initial position is quite satisfactory.

Table II shows the mean scores of duration of the [s] given by the experimental group and the control group. In both groups the duration of the [s] in the initial and final positions is adequate, but the duration of the [s] in clusters in the initial and final positions is much shorter than either native speakers.

TABLE II Duration (in ms) of the full mean frication of the test stimuli

Group	Experimental Group A			Control Group B		
	Pre	Post	+/-	Pre	Post	+/-
(single word)						
init.	see	150 258	100	163 178	15	
	sandwich	105 266	121	129 165	36	
	sixteen	120 290	70	121 105	-16	
mid.	assignment	96 325	229	93 157	64	
final	pass	275 423	148	270 285	15	
	face	198 488	290	188 320	132	
	bus	132 549	417	200 236	36	
	base	228 320	92	264 302	38	
	miss	260 383	123	266 316	50	
/s/ clus.						
init.	special	88 232	144	63 85	22	
	standard	66 240	174	94 82	-12	
mid.	sixteen	99 117	18	85 83	-2	
	telescope	93 160	67	117 120	3	
	textbook	104 157	53	140 135	5	
final	ask	50 140	90	49 82	33	
	best	65 215	150	73 123	50	
(words in sentence)						
init.	see	180 162	-18	156 160	4	
	sister	148 224	76	140 150	10	
	successful	139 162	23	121 134	113	
mid.	businessman	70 150	80	65 93	28	
final	nice	93 318	225	96 120	24	
	nurse	183 377	194	137 183	46	
/s/ clus.						
	successful	85 154	69	88 115	27	
	successful	79 321	242	99 115	16	
	sister	60 178	118	88 90	2	
Average			132	30		

In both groups, the duration of the speakers is about the same, (e.g. some duration is less than 100 ms, and some is more than 100 ms). Even though the range of duration does not vary between either group A or B of the native speakers, it is true that the native speaker's [s] sound is much sharper and clearer than that of the Japanese students because of their different breath stream energy levels. The ideal way of teaching a sharp [s] to the Japanese students is to train them to emit a strong breath. However, if the students cannot perform this, they should try to prolong the [s]. Obviously the longer the students prolong the [s], the longer the period the listener will hear the [s]. Although the prolonged [s] is not a sharp sound, the listener may be able to perceive the [s] adequately. Jongman (1989) discovered that the minimum frication duration required for identification varies from fricative to fricative and approximately an initial 50 ms or more is required for [s][3]. Longer duration is required for the speaker to remember the place of articulation for fricatives. However, if the duration of the [s] sound produced by the Japanese speaker with a weaker breath is 50 ms, the duration is insufficient for the English speakers to perceive the [s]. Therefore, for the Japanese speakers prolonging the [s] is important. The adequate way for the Japanese speakers to produce a sharp [s] sound is to produce it with the strong breath; however, if it is difficult for them to emit a strong breath stream, they should prolong the [s].

From Table II the length of the mean duration of the [s] produced by Group A increased more than four times higher than that of Group B. The duration of frication noise of the [s] clusters produced by native speakers is very long, while the Japanese students' duration is very short.

IV. SUMMARY

This study has examined the phonetic method of teaching Japanese university students. The phonetic method has been

used by many language teachers for teaching pronunciation. The experiment here described gave some indication that phonetic explanations and the provision of a few alternative steps of articulation bring a desirable result and are effective. This is viewed as evidence that detailed articulatory information, which allows university students to grasp the causes of their inappropriate pronunciation, will bring satisfactory results.

Further experiments need to be conducted to determine the criteria for evaluating the English voiceless fricative /s/ sound produced by the Japanese speaker in the language laboratory.

References

- [1] Fisher, Hilda B., *Improving Voice and Articulation*, Boston: Houghton Mifflin Comp., 1975.
- [2] Hughes, G. W., and Halle, M., "Spectral properties of fricative consonants," *J. Acoust. Soc. Am.* 28, 1956, 303-310.
- [3] Jongman, Allard, "Duration of frication noise required for identification of English fricatives," *J. Acoust. Soc. Am.*, 85 (4), 1989, 1718-1725.
- [4] Murakawa, H., *PROTS (Pronunciation Training System)*, Vol.3, Hamamatsu: Kawai Musical Instruments Manufacturing Co., 1989.
- [5] Whittaker, James O., *Introduction to Psychology*, Philadelphia: W. B. Saunders Company, 1966.
- [6] You, H. Y., "An acoustical and perceptual study of English fricatives," M.A. thesis, University of Edmonton, Canada, 1979.