# Contrast and covert contrast: The phonetic development of voiceless sibilant fricatives in English and Japanese toddlers

Fangfang Li[a,*], Jan Edwards[b], Mary E. Beckman[a]

[a]*Ohio State University, 222 Oxley Hall, 1712 Neil Avenue, Columbus, OH 43210, USA*
[b]*University of Wisconsin-Madison, USA*

## Abstract

This paper examines the acoustic characteristics of voiceless sibilant fricatives in English- and Japanese-speaking adults and the acquisition of contrasts involving these sounds in 2- and 3-year-old children. Both English and Japanese have a two-way contrast between an alveolar fricative (/s/), and a post-alveolar fricative (/ʃ/ in English and /ç/ in Japanese). Acoustic analysis of the adult productions revealed cross-linguistic differences in what acoustic parameters were used to differentiate the two fricatives in the two languages and in how well the two fricatives were differentiated by the acoustic parameters that were investigated. For the children's data, the transcription results showed that English-speaking children generally produced the alveolar fricative more accurately than the post-alveolar one, whereas the opposite was true for Japanese-speaking children. In addition, acoustic analysis revealed the presence of covert contrast in the productions of some English-speaking and some Japanese-speaking children. The different development patterns are discussed in terms of the differences in the fine phonetic detail of the contrast in the two languages.
© 2008 Elsevier Ltd. All rights reserved.

## 1. Introduction

In the first few years of life, children learn to produce most of the sounds of their native language, so that they are able to convey phonological contrasts using forms that are close enough to adult expectations to be regarded as "correct" by adult listeners. When compared across languages, this process of phonological development shows both striking similarities and clear-cut differences. Across many languages, we find that most children can produce the vowels in the ambient language by about age 2 and that stop consonants and glides develop relatively early, while fricatives, affricates, and liquids tend to be later-acquired (e.g., Dinnsen, 1992; Kent, 1992). Jakobson (1941/1968) suggested that the similarities observed across languages were due to universal substantive principles—"implicational laws"—that structure the phoneme inventories of all spoken languages and that also determine how children acquire speech sounds. Subsequent researchers

(e.g., Dinnsen, 1992; Ingram, 1989; Kent, 1992) have expanded on Jakobson's claim by suggesting how many of these cross-language similarities in phoneme acquisition might be attributed to such factors as phonetic constraints on production and perception.

One of Jakobson's implicational universals was that languages will have fricatives only if they also have stop consonants and that children would acquire fricatives only after they had acquired stop consonants. As Jakobson predicted, fricatives are less frequent than stops in the world's languages and fricatives also tend to be acquired later than stops in languages in which phonological acquisition has been studied. For example, the UPSID-PC database (Maddieson & Precoda, 1990) lists 38 languages that have no fricatives whatsoever, but no languages that do not have any stops. Furthermore, in English, both /d/ and /t/ are produced correctly in word-initial position by more than 90% of children by age 3, while word-initial /s/ is not produced correctly by that many children until after age 7 (Smit, Hand, Freilinger, Bernthal & Bird, 1990). A number of researchers have suggested that the greater production demands for fricatives relative to stops

*Corresponding author. Tel.: +1 614 327 0299; fax: +1 614 292 8833.
*E-mail address:* cathy@ling.ohio-state.edu (F. Li).

accounts for their later acquisition. As Kent (1992) points out, stops require only a more-or-less accurate ballistic gesture, while fricatives require a careful tongue body posturing that directs a high velocity of air at a very precisely located narrow constriction.

This paper compares the mastery of sibilant fricatives across two languages which contrast a more anterior dental or alveolar fricative with a more posterior one. Both English and Japanese have two sibilant fricatives. These two fricatives are normally assigned the phonological labels /s/ and /ʃ/ in English and either the same labels or the labels /s/ and /ɕ/ in Japanese. The second set of phonological labels for Japanese reflects the fact that, while the Japanese /ɕ/ is readily assimilated to English /ʃ/ by English speakers (as in *sushi*) and the English /ʃ/ is assimilated to Japanese /ɕ/ by Japanese speakers (as in /ɕerii/ *sherry*), the two post-alveolar fricatives have different articulatory configurations. In English, the coronal post-alveolar /ʃ/ contrasts with /s/ in tongue position; the constriction for /ʃ/ is further back in the oral cavity than for /s/ (Tabain, 2001). Furthermore, along with the more retracted position of the tongue in /ʃ/, it is also characterized as having a wider tongue groove and a resulting larger cross-sectional area (Fletcher & Newman, 1991; Gafos, 1999; Stone, Faber, Raphael, & Shawker, 1992). The contrast in Japanese is a contrast more of tongue posture than of place of articulation. The front of the tongue body is bunched up towards the palate to produce /ɕ/, but not /s/ (Akamatsu, 1997; Toda & Honda, 2003).

The acquisition of sibilant fricatives is protracted in both English and Japanese. For example, a large cross-sectional study of American-English-speaking children by Smit et al. (1990) found that at age 3, only 56% of children correctly produced word-initial /ʃ/ and only 62% correctly produced word-initial /s/. By contrast, 75% of English-speaking children correctly produced word-initial /f/ and over 90% correctly produced initial /d/ and /t/. Similarly, in Japanese, a cross-sectional study of Japanese 3-year-olds by Yasuda (1970) found that only 60% correctly produced initial /ɕ/ and only 25% of correctly produced /s/, and a larger norming study by Nakanishi, Owada, and Fujita (1972) reports errors for /s/ still in 6-year-olds. While lingual fricatives are late-acquired in both English and Japanese, the substitution patterns are different in the two languages. In English, the typical error pattern is the "fronting" of /ʃ/ to a perceived [s] substitution (e.g., Li & Edwards, 2006; Weismer & Elbert, 1982). Exactly the opposite error pattern is observed in Japanese-acquiring children, who are more likely to be perceived as substituting the more posterior [ɕ] for target /s/ (Nakanishi et al., 1972).

When acquisition of a contrast is protracted in this way, children may go through a stage of "covert contrast"—production of a perceptually unreliable, but statistically significant acoustic difference between two sounds. Covert contrast has been observed for a variety of contrasts,

including the voicing contrast for stop consonants (e.g., Macken & Barton, 1980; Maxwell & Weismer, 1982; Scobbie, Gibbon, Hardcastle, & Fletcher, 2000) and stop place of articulation (Forrest, Weismer, Hodge, Dinnsen, & Elbert, 1990; White, 2001). Many studies of covert contrast have focused on children with phonological disorders (see Scobbie (1998) for a review of covert contrast types), and it has been shown to be of clinical significance in that children who produce covert contrast have a better prognosis than children who produce no contrast at all (Tyler, Figurski, & Langdale, 1993).

There has been relatively little work on covert contrast in the acquisition of fricatives. Baum and McNutt (1990) observed covert contrasts in both amplitude and spectral shape between misarticulated /s/ (which was perceived as [θ]) and the target /θ/. Tsurutani (2004) also found some evidence of covert contrast in the productions by Japanese-acquiring children of a small number of words that happened to exemplify the /s/-/ɕ/ contrast in a larger study of other contrasts.

By definition, the study of covert contrast requires the use of instrumental measures in addition to a transcription analysis. For example, studies of covert contrast for stop voicing differences have measured voice onset time in word-initial consonant productions (e.g., Macken & Barton, 1980) and preceding vowel duration in word-final consonants (e.g., Maxwell & Weismer, 1982). In this paper, we first report on the acoustic measures that we developed in order to differentiate between the two sibilant fricatives in each of the two languages, measures which we then apply to children's productions, to look for covert contrast.

Previous research on languages such as English that have a place-of-articulation contrast for sibilant fricatives has focused on differentiating the two fricatives based on the spectral properties of the frication noise (Behrens & Blumstein, 1988; Hughes & Halle, 1956). The [−anterior] fricative /ʃ/ has a longer front cavity than /s/ both because of its more posterior place of articulation and also because of its characteristic lip rounding preceding unrounded vowels. This difference in front cavity length results in more low-frequency energy for the /ʃ/ spectrum and more high-frequency energy for /s/ (Hughes & Halle, 1956; Stevens, 1998).

A commonly used method for examining the spectral properties of fricative noise is spectral moments analysis, in which the power spectrum is treated as a probability distribution so that the mathematical moments can be calculated (Forrest, Weismer, Milenkovic, & Dougall, 1988; Jongman, Wayland, & Wong, 2000; Shadle & Mair 1996). The first spectral moment (the mean or "centroid" frequency) works well to distinguish between /s/ and /ʃ/ in English. In spectra with only one prominent mode, the frequency of the first moment is negatively correlated with the length of the front resonating cavity, and thus roughly describes where the constriction is made relative to the length of the oral cavity. The second spectral moment (standard deviation) does not seem to be useful in

distinguishing between the two sibilant fricatives of English, and it is mainly used to differentiate between a flat diffuse spectral shape, as in /f/, and a peaky, compact distribution as in /s/. But it may help to distinguish Japanese /s/ and /ɕ/, since according to Akamatsu (1997), Japanese /s/ is less sibilant than /ɕ/, and is less sibilant than English /s/ as well. The third spectral moment (skewness) may also be useful for distinguishing between /s/ and /ʃ/ in English, as it is also correlated with a place-of-articulation distinction. In general, /ʃ/ should have a positive value, indicating a concentration of energy in the lower frequencies below the mean value, while /s/ should have a negative value, indicating a concentration of energy in the higher frequencies above the mean value. The fourth spectral moment (kurtosis) may be useful for distinguishing between fricatives with tongue posture differences, as these differences result in changes in the peakiness of the spectral shape. Specifically, the compact fricative /ɕ/ might have a more prominent focalization of energy around a single peak than the diffuse /s/ of Japanese, and hence a higher kurtosis value.

Spectral moments analysis has also been used to classify English-speaking children's fricative productions. Nittrouer (1995) used it to compare the productions of /s/ and /ʃ/ in children aged 3–7, and adults. She found more variability in the children's productions relative to those of adults. Nissen and Fox (2005) found that the first and third spectral moments worked well to classify productions of /s/ and /ʃ/ by 3- to 6-year-old children.

We suspected that spectral moments analysis alone might not be sufficient for differentiating between /s/ and /ɕ/ in Japanese, as the contrast in Japanese does not primarily involve place of articulation. In previous work on Japanese and Polish, both of which have a similar tongue posture distinction for sibilant fricatives, researchers have used the second formant ($F2$) frequency taken at the onset of the vowel following the fricative to index the length of the back cavity in the fricative (Funatsu, 1995; Halle & Stevens, 1997). The tongue-bunching used to produce /ɕ/ creates a long palatal channel and a consequently shorter back cavity, which should result in a higher $F2$ frequency at the onset of the following vowel, relative to the frequency after /s/. Onset $F2$ frequency has also been shown to be reliable in distinguishing between /s/ and /ʃ/ in English (Jongman et al., 2000; Nittrouer, Studdert-Kennedy, & McGowan, 1989). Therefore, we decided to include both spectral moments and onset $F2$ frequency in our analysis.

As noted above, the purpose of this paper was to develop acoustic measures that would differentiate between the two sibilant fricatives in both English and Japanese and then to use these measures to examine the acquisition of contrast and covert contrast for sibilant fricatives in the two languages. We made two predictions. One was that the fine phonetic detail of the contrast would be different in the two languages, given the different articulatory configurations of the post-alveolar fricative in the two languages.

The second was that we would observe covert contrast in the productions of at least some of the children, given the protracted period of acquisition of this contrast in both languages. This paper differs from previous work in three important respects. First, it examines acquisition of a contrast that has been little studied and is known to be late and frequently misarticulated. Second, it examines the acquisition of this contrast across two languages. Third, it examines covert contrast in a relatively large population of normal-developing children in the two language communities.

## 2. Methods

### 2.1. Participants

The child participants for both languages were 2-year-olds and 3-year-olds, approximately ten children for each of these two age-groups for each language. We also collected data from five adult native speakers for each language: three females and two males for English and four females and one male for Japanese. All children had normal speech and language, based on parent and teacher report, and had passed a hearing screening using otoacoustic emissions at 2000, 3000, 4000, and 5000 Hz. All adult participants were undergraduate or graduate students at Ohio State University with no reported history of speech, language, or hearing problems. Table 1 gives information on the age of the child participants and the exact number per age group. The English-speaking children and the adult speakers of both languages were tested in Columbus, OH. The Japanese-speaking children were tested in Tokyo and Hamamatsu, Japan. All children were monolingual speakers of their native language, while the Japanese-speaking adults had lived outside of Japan for less than 5 years.

### 2.2. Materials

The materials were word-initial voiceless fricatives, elicited in a word-repetition experiment that was designed to sample all of the lingual obstruents in each target language. The words were chosen so that the target

Table 1
Mean age in months (standard deviation in parentheses) and number of subjects for child participant groups for English and Japanese.

| Age groups: | Language | English | Japanese |
| --- | --- | --- | --- |
| 2-year-olds | Age | 31 (3.4) | 32 (1.8) |
|  | *N* | 9 | 13 |
| 3-year-olds | Ages | 39 (2.6) | 44 (2.4) |
|  | *N* | 12 | 9 |

obstruents would be sampled before a variety of following vowels, chosen from a set of five categories that we will call /i e a o u/. Japanese has only these five vowels. For English, we collapsed together vowels that have similar co-articulatory effects. Specifically, we included both lax and tense vowels in each vowel category where the tense/lax contrast is relevant (for example, both /i/ and /ɪ/ were included in the /i/ category) and we included all three low back vowels /ɑ, ɔ, ʌ/ in the /a/ category. We elicited these word-initial CV sequences in familiar pictureable words in both English and Japanese. There were approximately three target words for each CV sequence. Not all of the CV sequences could be elicited. This is because */si/ is unattested in Japanese and /çe/ is attested only marginally, primarily in recent loan words from languages such as English. In English, only two words containing /ʃu/ were elicited because there are few words containing this sequence that are familiar to young children. A complete list of the target words for each language is given in the appendix.

For both languages, the stimulus items for the word-repetition task were spoken by an adult female native speaker in a child-directed speech register. The speaker was familiar with the purpose of the task. The fricative-initial words were recorded in a mixed list along with words that began with other lingual obstruents (stops and affricates). The adult female speakers' productions were digitally recorded at a sampling rate of 22,500 Hz. For each word type, three tokens were presented to adults and then two tokens that were perceived with at least 80% accuracy by the five adult native speakers were selected for use with the children. In the presentation to the children, each word type was paired with a color photograph that was culturally appropriate for the particular language and country.

## 2.3. Procedure

Adults were tested in a quiet room individually. For the adults, each stimulus item was played out over speakers connected to a computer sound card and the adult participants were asked to repeat each item just as they heard it. Their responses were recorded on a Marantz CD recorder, using a high-quality head-mounted microphone.

In both countries, children were tested in a quiet room in a preschool individually. For the children, each trial item consisted of a picture and the associated sound file, which were presented simultaneously to the participant over a laptop with a 14-in screen using a program written specifically for our purposes. The computer program included an on-screen VU meter to help the children monitor their volume and a picture of a duck walking up a ladder on the left side of the screen to provide visual feedback to the children about how close they were to completing the task. There was a practice session prior to administration of the experimental task. The English-speaking children were instructed as follows: "You are

going to see some pictures on my computer and hear some words. Your job is to repeat the words you hear. So if you hear the computer say "ball," what are you going to say?" The instructions were similar in Japanese. The task was relatively simple and the children learned it easily. During the experiment, children were asked to repeat responses in the following cases: (1) if the response was different from the prompted word (e.g., the child said *duck* when prompted with *goose*) or (2) if the tester thought the target sequence would be impossible to transcribe because the response was spoken very softly, or overlapped with the prompt or with background noise (e.g., a door slam). The children's responses were recorded directly onto a CD or a digital audiotape, using a high-quality head-mounted microphone.

## 2.4. Transcription

All audible responses were transcribed and included in the statistical analyses. A native speaker/trained phonetician transcribed all initial target CV sequences, using both the audio signal and the acoustic waveform. The English data were transcribed by an American-English speaker and the Japanese data were transcribed by a Japanese speaker. Both transcribers were from the same dialect region as the child participants. The fricatives were transcribed as either correct or incorrect. The native speaker also transcribed substitution errors when the target consonant was categorized as incorrect. The transcriptions were based on the word-initial consonant-vowel sequence which the transcribers could isolate on the waveform and listen to as often as necessary. The transcribers transcribed on a child-by-child basis so that both sibilant fricatives (and all other target word-initial obstruent sounds) were transcribed for one child before moving on to the next. The transcribers always knew what the target word (and fricative) was. A second native speaker independently transcribed 20% of the data using the same methodology. Phoneme-by-phoneme inter-rater reliability was 90% for English and 89% for Japanese.

## 2.5. Acoustic analyses

Table 2 summarizes all of the acoustic measurements that were made. We used Praat (Boersma & Weenink, 2005) for all of the acoustic analyses. The onset of the fricative was defined as the first appearance of aperiodic noise on the waveform, simultaneously accompanied with frication noise above 2500 Hz from the spectrogram. The offset of the fricative was defined as the first zero-crossing of the periodic waveform of the following vowel.

For the spectral moments analysis, an FFT spectrum was made over a 40 ms Hamming window centered at the midpoint of the fricative noise. The middle 40 ms window was chosen because it is the most steady-state portion of the fricative noise and is least likely to be influenced by amplitude effects at the start up of the fricative or by

Table 2
Summary of parameters for acoustic analyses.

|  | Acoustic parameter | Definition | Articulatory interpretation |
|---|---|---|---|
| Fricative spectrum moments | M1 (centroid) | Center of mass of the distribution (the weighted mean frequency) | Negatively correlates with the length of the front resonating cavity |
|  | M2 (standard deviation) | Spread of the distribution (average squared distance from the centroid) | Differentiates tongue posture between apical and laminal |
|  | M3 (skewness) | Asymmetry in the spectral shape (the difference between the spectrum below the centroid and the spectrum above the centroid) | Negatively correlates with the length of the front resonating cavity |
|  | M4 (kurtosis) | Peakiness of the spectral shape (the average distance from the centroid raised to the fourth power, divided by the squared variance of the distribution) | Dfferentiates tongue posture between apical and laminal |
| CV transitions | Onset *F*2 frequency | *F*2 frequency at the onset of the following vowel | Negatively correlates with the length of the back resonating cavity |

anticipatory co-articulation with the vowel. The setting in Praat that we used to estimate the onset *F*2 was an LPC analysis specified for 5 formants (10 coefficients) calculated over a range from 0 to 5500 Hz for adults and from 0 to 7000 Hz for children. The window length was 0.025 ms. We hand-corrected mistracked *F*2 values for seven tokens in English and five tokens in Japanese. All calculations were made without pre-emphasis.

## 3. Results

### 3.1. Transcription

Transcription analysis of elicited single word productions has traditionally been used to describe the age at which most children correctly produce a particular consonant (e.g., Smit et al., 1990). These "developmental norms" for consonant mastery are used primarily to help with clinical diagnosis of speech sound disorders. Mastering a consonant means that a child is able to produce the sound in a form that adult listeners accept as correct. More specifically, the operational definition in the literature for "mastery" of a speech sound typically is 75% accuracy for an individual child in a particular word position (e.g., Prather, Hedrick, & Kern, 1975; Smit et al., 1990; Templin, 1957). Similarly, the criterion used for mastering the contrast between two sounds is 75% accuracy for both sounds in a particular word position. We adopted these operational definitions to determine how many of the English- and Japanese-speaking children had mastered each of the two fricatives and the contrast between them, as shown in Table 3. Two observations are of interest: first, /s/ is mastered by more children than /ʃ/ in English, while /ɕ/ is mastered by more children than /s/ in Japanese and this is especially true for the 2-year-old group. Further, more English-speaking children than Japanese-speaking children have mastered /s/ by age 3 ($\chi^2$ (1,22) = 6.5, $p < 0.01$). At the

Table 3
Number of children with 75% or more correct productions in each language, based on transcription analysis.

| Fricative | English (N = 22) | |
|---|---|---|
|  | 2-year-olds | 3-year-olds |
| /s/ | 9 | 8 |
| /ʃ/ | 2 | 6 |
| /s/-/ʃ/ contrast | 1 | 5 |
|  | Japanese (N = 21) | |
| /s/ | 0 | 5 |
| /ɕ/ | 2 | 5 |
| /s/-/ɕ/ contrast | 0 | 3 |

same time, although more English children appear to have acquired the contrasts between the two sibilant fricatives than Japanese-speaking children, the pattern is not statistically significant, and it has to await future studies by testing more children in this age range.

Table 4 shows the most frequent substitution processes for the two languages. In English, by far the most common process is "fronting" of /ʃ/ to [s], both in terms of the number of children who produced this substitution and in terms of the total number of substitutions made by all of the children. By contrast, in Japanese, the most common process is "backing" of /s/ to [ɕ], although this substitution pattern does not outnumber other substitution processes as much as the English error pattern does. In general, fronting errors predominate in English while backing errors predominate in Japanese. In both languages, no major vowel effect on the substitution patterns was observed. In English, /ʃ/-to-[s] substitutions occurred in front of all five vowels and at a similar rate (/a/: 20; /e/: 32; /i/: 23; /o/: 21;

Table 4
The most frequent substitution processes in the productions of English-speaking and Japanese-speaking children.

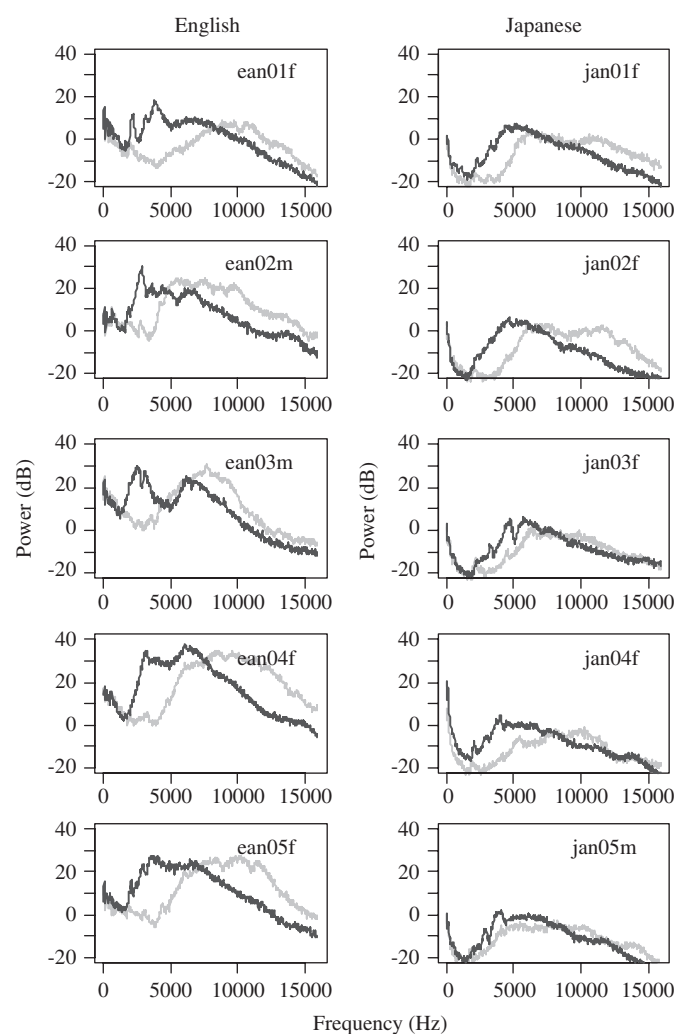| | Error pattern | English | | | Japanese | | |
|---|---|---|---|---|---|---|---|
| | | Error type | Num of children | Num of instances | Error type | Num. of children | Num. of instances |
| Place error | Fronting | /ʃ/-> [s] | 12 | 116 | /ç/ ->[s] | 7 | 18 |
| | | /s/ ->[f, v] | 13 | 39 | | | |
| | | /s/ -> [θ] | 2 | 2 | | | |
| | Backing | /s/-> [ʃ] | 2 | 6 | /s/-> [ç] | 11 | 63 |
| | | | | | /ç/ -> [ç] | 3 | 12 |
| Manner error | Stopping | /s/-> [tʰ, t] | 5 | 6 | /s/-> [t, d] | 8 | 25 |
| | | /ʃ/-> [tʰ] | 1 | 1 | /ç/-> [t, d] | 7 | 21 |
| | | Other | 3 | 9 | other | 1 | 1 |
| | Affrication | /s/-> [tʰs, ts] | 4 | 14 | /s/-> [ts, dz] | 2 | 3 |
| | | /ʃ/-> [tʃ] | 4 | 26 | /ç/-> [tç] | 7 | 31 |
| | | /s/-> [tʃ] | 1 | 1 | /s/-> [tç] | 7 | 19 |
| | | /ʃ/-> [tʰs] | 2 | 2 | | | |
| | Other | /ʃ/-> [h] | 1 | 1 | /ç, s/-> [h] | 2 | 2 |



Fig. 1. Averaged spectra of /s/ (gray) and /ʃ/ (or /ç/) (black) for productions of adult speakers of English (left) and Japanese (right).

/u/: 20). In Japanese, /s/-to-[ç] substitutions occurred in front of all four vowels where /s/ is attested, that is, before all vowels except /i/, which is a phonotactically illegal environment for /s/. Again, for Japanese, the substitution rate was similar across the rest of the four vowels (/a/: 14; /e/: 15; /o/: 16; /u/: 18).

### 3.2. Acoustic analyses—adult productions

Fig. 1 plots the averaged spectra of /s/ and /ʃ/ or /ç/ for all five speakers of English and of Japanese, respectively. Despite the inter-speaker differences in the shapes of the spectra, these five speakers demonstrate great consistency in the overall patterning that contrasts /s/ with /ʃ/ or with /ç/ in the two languages. It can be observed that the average spectra of the two sibilant fricatives have more distinct patterns from each other in the English-speakers' productions, as compared to those of the Japanese speakers. Moreover, the spectrum for /ç/ in Japanese is consistently peakier than the spectrum for /s/ in Japanese.

Fig. 2 plots the mean of the five acoustic parameters for /s/ and /ʃ/ or /ç/ for each vowel context for both languages for female speakers of the two languages (the male speakers' productions pattern similarly). Note that the first spectral moment (centroid) effectively captures the differences in spectral energy concentration, yielding higher values for /s/ than for /ʃ/ or /ç/. Higher centroid values indicate a higher-frequency energy concentration for /s/, and thus a more front lingual constriction for /s/ as compared to /ʃ/ or /ç/. The two voiceless sibilant fricatives are better separated by the centroid values in English as compared to Japanese, which is in accordance with the observations from the averaged spectra for the two languages shown in Fig. 1. The second spectral moment (standard deviation) is larger for /s/ than for /ʃ/ or /ç/,
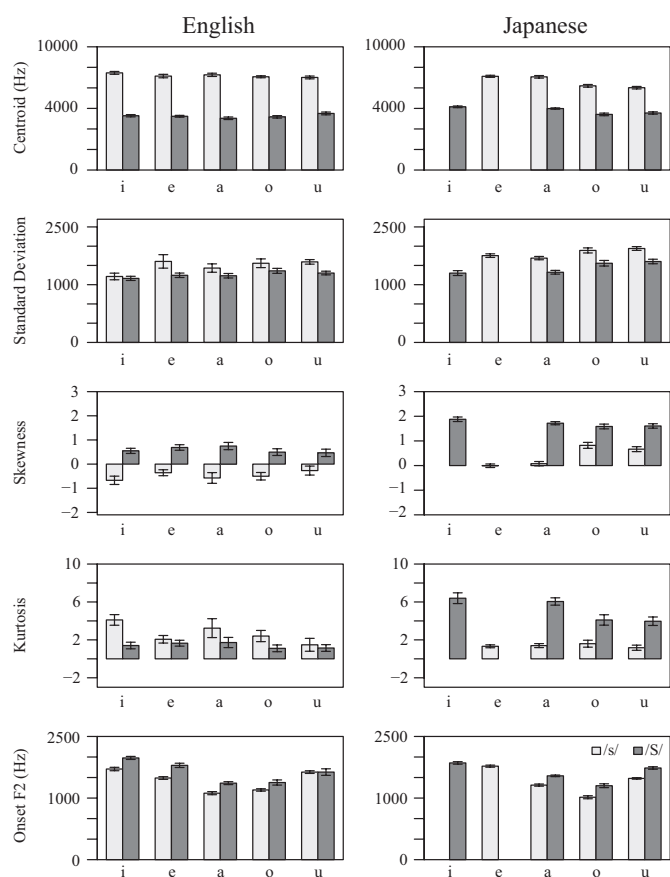
Fig. 2. Means of first four spectral moments and *F*2 onset frequency by vowel context for productions of /s/ and /ʃ/ (or /ɕ/) by English-speaking (left) and Japanese-speaking (right) female adults. /s/ in the legend represents the alveolar/dental fricative in both English and Japanese, and /S/ represents the post-alveolar fricatives (/ʃ/ in English and /ɕ/ in Japanese). This convention is used in all subsequent graphs. The error bars represent the standard error about the mean.

Table 5
*F*-values from two-way ANOVA analyses on each of the five acoustic parameters for productions of adult speakers; English speakers in top five rows, Japanese speakers in bottom five rows.

| Talker | Fricative spectrum | | | | CV transition |
|---|---|---|---|---|---|
| | *M*1 | *M*2 | *M*3 | *M*4 | Onset *F*2 |
| ean01f | 767.6*** | 16.0*** | 214.6*** | 5.6* | 43.3*** |
| ean02m | 446.4*** | | 48.5*** | 7.6** | 92.1*** |
| ean03m | 1215.3*** | 188.0*** | 266.1*** | 165.0*** | 28.1*** |
| ean04f | 975.3*** | 10.4** | 6.8* | 29.0*** | 22.3*** |
| ean05f | 1479.2*** | | 110.2*** | 0.3* | 11.2** |
| jan01f | 272.4*** | 46.1*** | 56.2*** | 49.0*** | 13.0*** |
| jan02f | 1018.4*** | 93.3*** | 226.6*** | 128.1*** | 30.4*** |
| jan03f | 258.8*** | 8.6** | 189.3*** | 77.2*** | 104.6*** |
| jan04f | 700.8*** | 10.4** | 213.5*** | 14.8*** | 81.4*** |
| jan05m | 374.5*** | 83.8*** | 206.7*** | 62.4*** | 59.0*** |

Empty cells indicate a non-significant main effect of fricative type.
***$p < 0.001$; **$p < 0.01$; *$p < 0.05$.

indicating a more diffuse shape for /s/ in the two languages. The onset *F*2 frequency also shows similar patterns for /s/ as compared to /ʃ/ or /ɕ/ for both languages. The most notable difference between the two languages is for the fourth spectral moment (kurtosis). In English, the post-alveolar /ʃ/ has a smaller kurtosis value than /s/, while the opposite pattern is observed in Japanese. The particularly high kurtosis value for Japanese /ɕ/ reflects the more compact and symmetrical distribution of energy around a single peak seen in the averaged spectra in Fig. 1.

We performed five two-way ANOVAs for each adult speaker. The within-subject factors were fricative and vowel, and the dependent variables were the five acoustic measures. Table 5 shows the results of these analyses. It can be observed that the *F*-values for the centroid frequency (first spectral moment) are generally an order of magnitude greater than the *F*-values for the other acoustic parameters, suggesting that the first spectral moment is the primary acoustic parameter for distinguishing between the two fricatives in both languages.

We used logistic regression in a hierarchical linear model (HLM; Raudenbush, Bryk, Cheong, & Congdon, 2004) to determine which acoustic parameters are needed to predict the categories of the two voiceless sibilant fricatives in each language. In each of these models, the dependent variable was the target fricative category (coded as 0 for /s/ versus 1 for /ʃ/ or /ɕ/) and the independent variables were *z*-score normalized values for each of the five acoustic parameters nested within subject, which was specified as a random grouping variable. This hierarchy of individual tokens nested within subjects allows us to evaluate effects of the different acoustic parameters simultaneously at the level of individual production tokens and at the level of the subjects, so as to control for the non-independence of tokens produced by the same speaker without ignoring within-subject variability. For English, once inter-subject differences were controlled for in this way, the first spectral moment alone perfectly classified the two fricatives. In Japanese (see Table 6), a combination of the first spectral moment and onset *F*2 frequency was needed to categorize the two fricatives. That is, as Table 6 shows, the coefficients for these two factors are significantly different from zero. Also, the coefficient for the first spectral moment is negative whereas that for the onset *F*2 value is positive, indicating that the log odds of the fricative being /ɕ/ is higher for lower centroid values and higher *F*2 onset values. Fig. 3 illustrates the fricative productions by English speakers and Japanese speakers, as described by centroid frequency and onset *F*2 frequency. It is clear that the English /s/ and /ʃ/ are clearly separated in the centroid dimension, but not in the onset *F*2 dimension. The Japanese /s–ɕ/ contrast, however, can be separated in the centroid dimension, although the separation is not as clear as that in the English speakers. At the same time, Japanese /s/ and /ɕ/ also differ from each other in their distributions in the onset *F*2 dimension.

## 3.3. Acoustic analyses—children's productions

We used the same acoustic analyses for the productions of all of the children who produced a contrast between the two sibilant fricatives or who produced [s] for /ʃ/ (in English) or [ç] for /s/ (in Japanese) substitutions. We then performed the same set of five two-way ANOVAs for these children's productions in the two languages so that we could compare the results of the transcription analysis to the results of the acoustic analysis and also

so that we could identify instances of covert contrast (see Table 7).

The rows of Table 7 with talker ID in italics give the results of the ANOVAs for the six English-speaking and the three Japanese-speaking children who were transcribed as producing a contrast between the alveolar and post-alveolar fricative. Two observations can be made. First, there is a significant difference in centroid frequency for the productions of all children from both languages who were transcribed as producing a contrast between the two sibilant fricatives. Second, the two fricatives are less well separated in the productions of the children, as compared to those of the adults. This can also be observed in Fig. 4, which plots onset $F2$ frequency against the first spectral moment for a representative English-speaking adult and an English-speaking child who produced a contrast between the two fricatives in both the transcription and the acoustic analysis.

Table 7 also gives the results of the ANOVAs for the children who were consistently transcribed as producing the most typical substitution errors—i.e., [s] for /ʃ/ in English or [ç] for /s/ in Japanese—but who nonetheless produced a statistically significant difference in at least one dimension. These are the children whose talker IDs are in boldface. They comprised 4 of the 16 English-speaking

Table 6
Results of hierarchical linear model for Japanese adult productions of /s/ and /ç/.

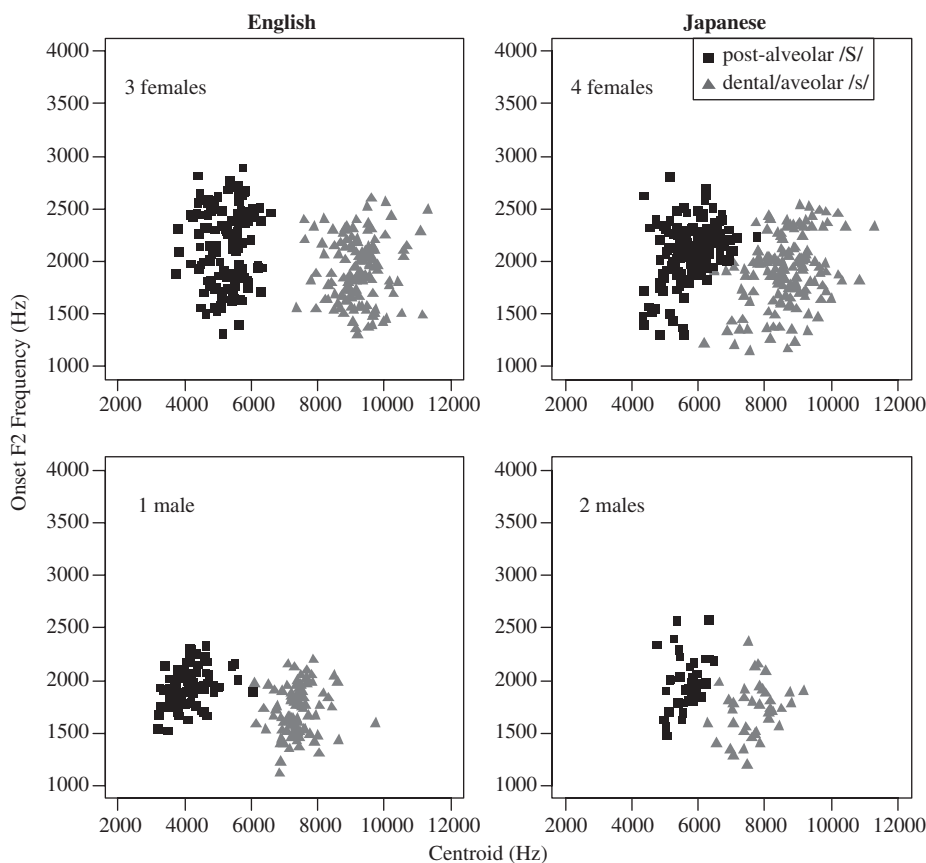| Coefficient for | Estimate | Standard error | $t$-Value | df | $p$-Value |
|---|---|---|---|---|---|
| Intercept | −1.571 | 1.735 | −0.905 | 4 | 0.417 |
| Centroid ($M1$) | −10.255 | 1.995 | −5.139 | 378 | **<0.001** |
| Standard Deviation ($M2$) | −1.694 | 0.934 | −1.813 | 378 | 0.070 |
| Skewness ($M3$) | 1.109 | 1.125 | 0.985 | 378 | 0.325 |
| Kurtosis ($M4$) | −1.626 | 1.195 | −1.361 | 378 | 0.174 |
| Onset $F2$ | 1.711 | 0.706 | 2.425 | 378 | **0.016** |

Significant $p$-values are in bold.



Fig. 3. Onset $F2$ frequency plotted against centroid frequency for the sibilant fricative productions of English-speaking adults (left) and Japanese-speaking adults (right).

children and 2 of the 18 Japanese-speaking children who did not have an overt contrast by our criterion of producing 75% of both fricatives correctly. We classified these productions as showing covert contrast. The *F*-values were smaller for the children with covert contrasts, as compared to the *F*-values for the children with overt contrasts, suggesting that covert contrast tends to be less stable and more variable. A second finding of interest is that the covert contrasts for the productions of all but one

Table 7
*F*-values from two-way ANOVAs on all five acoustic parameters for children with contrast (italic) or covert contrast (boldface), with English-speaking children's productions in first 10 rows, and Japanese-speaking children's productions in bottom 5 rows.

| Talker | Fricative spectrum | | | | CV transition |
|---|---|---|---|---|---|
| | *M*1 | *M*2 | *M*3 | *M*4 | Onset *F*2 |
| *e2n10m* | 79.0*** | 27.2*** | 39.1*** | 11.9** | 16.4*** |
| *e3n00f* | 61.6*** | 6.9* | 18.8*** | | 21.8*** |
| *e3n01m* | 62.3*** | 11.0* | | | |
| *e3n03f* | 238.9*** | 10.6** | 35.8*** | | |
| *e3n05f* | 121.6*** | | 7.6* | | 20.2*** |
| *e3n11f* | 205.0** | | 13.2*** | | 12.3** |
| **e2n01m** | | | | | 8.8* |
| **e2n03m** | 11.2* | | | | |
| **e3n07m** | | | | 4.7* | |
| **e3n12m** | | | 5.1* | | |
| *j3n01m* | 14.4** | 86.7*** | 28.9*** | 19.7*** | |
| *j3n09m* | 54.0*** | 75.3*** | 36.2*** | 13.9** | 176.0*** |
| *j3n12f* | 14.9** | | 5.5* | | |
| **j2n14f** | | | | 9.8* | |
| **j3n15m** | | | | | 10.1** |

Blank cells indicate no statistical significance.
\*\*\*$p < 0.001$; \*\*$p < 0.01$; \*$p < 0.05$.

of these children were identified by significant differences in a parameter other than the first spectral moment.

Examples of covert contrast are shown in Figs. 5 and 6. Fig. 5 plots the average spectra of target /s/ and /ɕ/ for two Japanese-speaking children. One of these children (j3n01m) was transcribed as having acquired the contrast between the two sibilant fricatives and his productions showed a significant difference between /s/ and /ɕ/ for all four spectral moments. For the other child (j2n14f), productions of both fricatives were transcribed as [ɕ]. The acoustic analysis of her productions revealed a significant difference between the two fricatives only in kurtosis (the fourth spectral moment). The averaged spectra of the [ɕ]-for-/s/ substitutions and that of the target /ɕ/ productions are similar, except that the [ɕ]-for-/s/ substitutions have a flatter energy distribution in the spectrum than the target /ɕ/ productions.

A different pattern of covert contrast was observed for one English-speaking child (e2n01m). The productions of this child showed evidence of a contrast between the two fricatives in onset *F*2 frequency. Fig. 6 plots onset *F*2 frequency against centroid frequency for the productions of this child as compared to the productions of a child transcribed as having a clearly acquired contrast. It can be observed that the productions of this child with a covert contrast separate out the two categories roughly in the dimension of onset *F*2 frequency, with no separation in centroid frequency, whereas the productions of the child with a clear contrast make a distinction more in the dimension of centroid frequency.

## 4. Discussion and conclusion

Three findings were of note in this study. First, we observed differences between English and Japanese for the two voiceless sibilant fricatives in the productions of adult native speakers. Differences were observed even for /s/, the
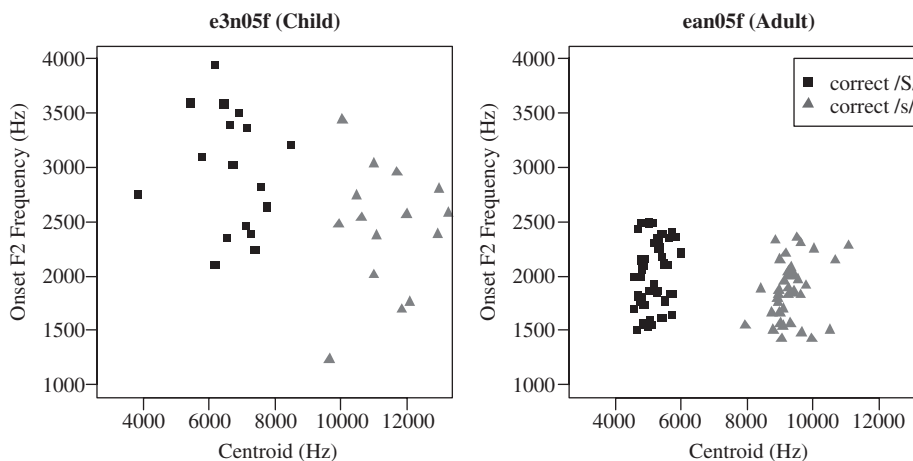


Fig. 4. Onset *F*2 frequency plotted against centroid frequency for an English-speaking child with a clear contrast between /s/ and /ʃ/ (left) and an English-speaking adult (right).
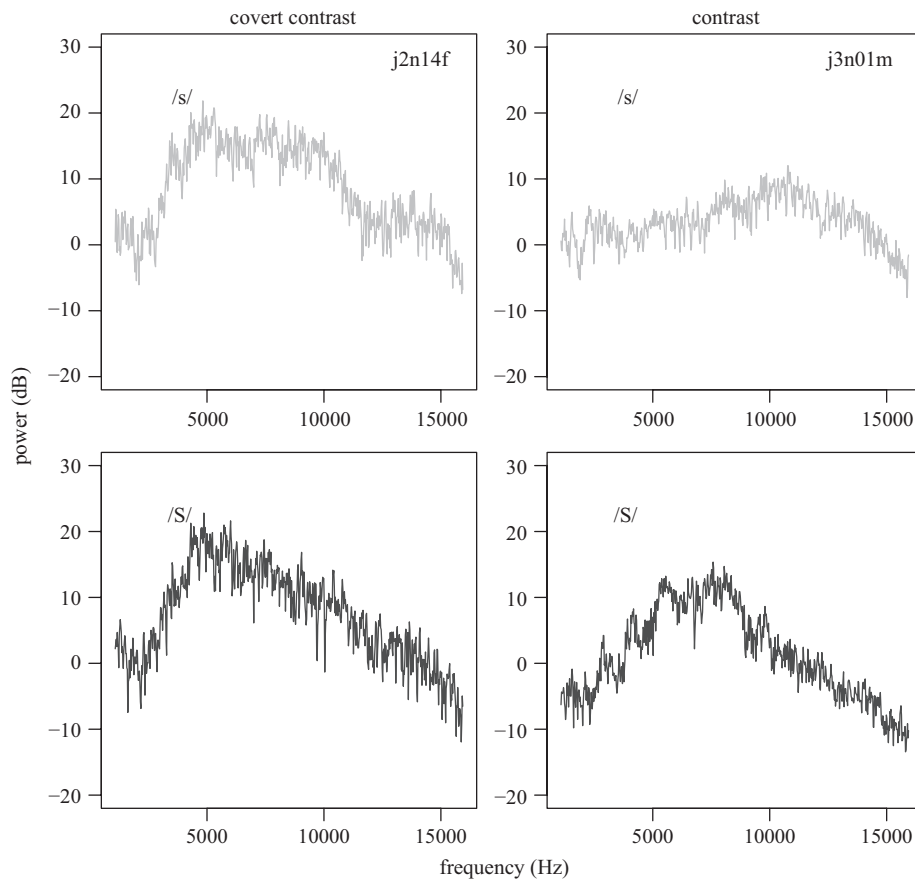
Fig. 5. Averaged spectra of target /s/ and /ɕ/ from the productions of a Japanese-speaking child with a covert contrast (left) and a Japanese-speaking child with a clear contrast (right).
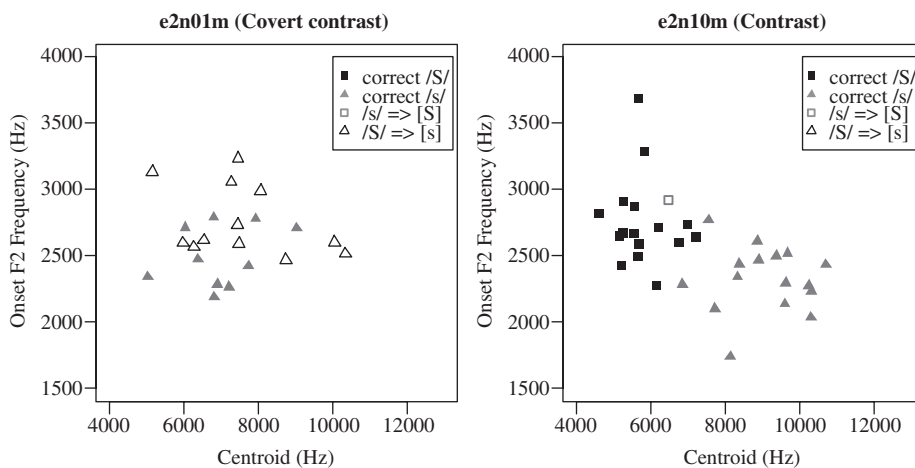
Fig. 6. Onset $F2$ frequency plotted against centroid frequency for an English-speaking child with a covert contrast (left) and an English-speaking child with a clear contrast (right).

fricative which the two languages have in common. Second, we also observed language-specific patterns in the productions of English-acquiring and Japanese-acquiring children. Finally, we found acoustic evidence of covert contrast between the two sibilant fricatives in the produc-

tions of some English- and Japanese-speaking children who did not appear to have a contrast in the transcription analysis. The cross-language differences in the adults' productions involved both the parameters used and the degree of separation observed. English /s/ and /ʃ/ were

clearly separated in terms of just a single acoustic parameter, the first spectral moment of the fricative noise or centroid frequency. By contrast, Japanese /s/ and /ɕ/ were less clearly separated in the centroid dimension, and centroid frequency has to be combined with onset *F*2 frequency in order to completely separate the two fricatives. We also found that /s/ has a more diffuse spectral distribution than its post-alveolar counterpart, /ɕ/, in Japanese (see Fig. 1), whereas there is no consistent pattern with respect to peakiness between the two sibilant fricatives in English. This can also be observed in Fig. 2, where the value of kurtosis (a measure of peakiness) for the Japanese /s/ is much lower than that of the Japanese /ɕ/. The value of kurtosis for the Japanese /s/ also tends to be lower than that of the English /s/. At the same time, the value of the second moment, the standard deviation, of the Japanese /s/ is generally higher than that of the English /s/. Taken together, these patterns suggest that the Japanese /s/ is more laminal and therefore has a more diffuse and less peaky spectral distribution than the English /s/. That is, these different acoustic realizations may result from subtle differences in articulatory configurations for a possibly more apical and clearly alveolar /s/ in English versus a more laminal and possibly somewhat dentalized /s/ Japanese. Therefore, learning to produce /s/ in English may not be the same as learning to produce /s/ in Japanese.

The cross-language differences that we saw in the children's productions involved both the transcription analyses and the acoustic analyses. More English-speaking children were transcribed as correctly producing /s/ than /ʃ/, while more Japanese-speaking children were transcribed as correctly producing /ɕ/ than /s/. Similarly, the most common error pattern for English-speaking children was fronting errors ([s] for /ʃ/ and [θ] for /s/ substitutions), while the most common error pattern for Japanese-speaking children was backing errors ([ɕ] for /s/ and [ç] for /ɕ/ substitutions). Locke (1983) claims that fronting is universal in the acquisition of fricatives across languages. Our results are consistent with earlier results of single-language studies that report opposite error patterns in fricative acquisition in English and Japanese. Like these earlier studies, we show that it is an over-simplification to describe children's acquisition just in terms of fronting or backing.

It is also interesting to note that /s/ was produced correctly by so many of the relatively young English-speaking children (2- and 3-year-olds). This was an unexpected finding, given that production of /s/ requires fine force regulation of frication (e.g., Kent, 1992). The results of Li (2008; also see Munson et al., 2008) suggest that this relatively early acquisition of /s/ in English is related to adults' perceptions as well as to children's productions. More specifically, Li and colleagues found that naïve adult English listeners accept a wider range of centroid frequencies for correct /s/ than their Japanese counterparts.

At the same time, one similarity across languages that we observed was that there was a significant difference in centroid frequency for the productions of all children who were perceived as producing a contrast between the two sibilant fricatives. Moreover, among the five parameters tested, the centroid frequency seemed to be the primary correlate of the contrast for both languages in the productions of both children and adults. In fact, it has been shown in a number of perception studies that spectral characteristics override transitional cues in the perception of voiceless sibilant fricatives, especially for English (LaRiviere, 1975; Whalen, 1984; Fernandez, Feijoo, Balsa, & Barros, 2000). In these experiments, the spectral characteristic that was manipulated was the frequency pole, a manipulation that produces differences of the sort that were measured by the centroid frequency in the current study.

Another similarity we found was that covert contrast was observed in both languages. We had predicted this result because of the protracted development of the contrast between the two sibilant fricatives in both English and Japanese. Out of 22 English-speaking children, only 6 showed complete mastery of the contrast at a level of 75% accuracy or more for both sounds, 4 children showed covert contrast as evidenced by instrumental analysis, and 12 children showed no indication of contrast mastery either in the impressionistic transcription or in the acoustic analysis. Out of 21 Japanese-speaking children, only three had mastered this contrast by age 3, two showed covert contrast, and 16 did not show any mastery of the contrast. We observed two forms of covert contrast. Most of the children with covert contrast used a non-primary parameter to differentiate between the two sibilant fricatives. One child with covert contrast differentiated the two fricatives with the primary parameter (centroid), but the difference was not large or consistent enough to be recognized by adults. While covert contrast was observed in both languages, however, it is important to note that the direction of the emerging contrast was different. The English-acquiring children with covert contrast were just beginning to distinguish acoustically between target /s/ and transcribed [s]-for-/ʃ/ substitutions, whereas the Japanese-acquiring children with covert contrast were just beginning to distinguish acoustically between target /ɕ/ and transcribed [ɕ]-for-/s/ substitutions. It should be noted that the only error patterns that we examined for covert contrast were [s]-for-/ʃ/ substitutions in English and [ɕ]-for-/s/ substitutions in Japanese. Other error patterns were also observed (although less frequently) such as [θ]-for-/ʃ/ substitutions in English and, potentially, children may produce covert contrast on these patterns as well.

A question for future research is to determine whether these opposite error patterns can be attributed solely to the cross-linguistic phonetic differences between the two languages. While the acoustic analyses revealed distinctive differences in the error patterns in the two languages, it is also possible that these different patterns might be related,

in part, to cross-linguistic perceptual differences that could be related to the different distribution of fricatives in the two languages. If English-speaking children made backing errors, it might be difficult for a native English-speaking transcriber to categorize them, since there are no English fricatives that have a more back place of articulation than /ʃ/. Similarly, if a Japanese-speaking child produced fronting errors, these would be difficult for a native Japanese-speaking transcriber to categorize since Japanese does not have any fricatives that are more front than /s/. Thus, it may be the case that these error patterns are not as different as they appear. Until we have judgments of the English-speaking children's productions by Japanese listeners and of the Japanese-speaking children's productions by English listeners, we should be cautious in concluding that the children in the two languages produce different substitution patterns.

In sum, we found that the fine phonetic detail of the two-way contrast in lingual sibilant fricatives differs considerably between English and Japanese. These language-specific differences affect acquisition of these sounds, as judged by an experienced native-speaker transcriber. More English-speaking 2- and 3-year-old children had mastered /s/, as compared to Japanese-speaking children of the same age. We have suggested that these language-specific differences in acquisition are related to differences in how the fricative contrast is represented acoustically between English and Japanese, as well as to the different distributional patterns in phonological representations between these two languages. Covert contrast was also observed in both languages. Four English-speaking and two Japanese-speaking children showed a significant difference between the two sibilant fricatives in one of the measured acoustic parameters in spite of the fact that the experienced native-speaker transcriber had transcribed all productions as /s/ (for English) or as /ɕ/ (for Japanese). The acoustic measures revealed cross-linguistic differences in /s/ as well as the presence of convert contrast. These results suggest that transcription alone is not adequate to describe phonological acquisition, since it is filtering children's production through adults' perceptual norms. Acoustic analysis is a useful tool in objectively describing children's productions unbiased by adults' perception. Perception experiments with naïve listeners can also help to objectively describe adults' perceptual norms in relation to these acoustic measures. These perceptual experiments are now underway.

## Acknowledgments

## Appendix A

Table A1 shows the stimuli for the two languages.

Table A1
Stimuli for the two languages.

| Target CV | English | | Japanese | |
|---|---|---|---|---|
| | Transcription | Orthography | Transcription | Gloss |
| si | /sit/ | *seat* | | |
| | /sikrɪt/ | *secret* | | |
| | /sɪk/ | *sick* | | |
| | /sɪstɚ/ | *sister* | | |
| se | /sef/ | *safe*[1] | /semi/ | 'cicada' |
| | /sel/ | *sail* | /senaka/ | 'back' |
| | /sem/ | *same* | /seNse:/ | 'teacher' |
| | /se/ | *say* | | |
| sa | /sɑ/ | *saw* | /saru/ | 'monkey' |
| | /sɑkɚ/ | *soccer* | /sakana/ | 'fish' |
| | /sɑs/ | *sauce* | /sakura/ | 'cherry blossom' |
| | /sɑk/ | *sock* | | |
| so | /sofə/ | sofa | /sori/ | 'slide' |
| | /sodə/ | *soda* | /sora/ | 'sky' |
| | /sop/ | *soap* | /so:se:dʒi/ | 'sausage' |
| su | /sup/ | *soup* | /sudzume/ | 'sparrow' |
| | /sutkes/ | *suitcase* | /suika/ | 'watermelon' |
| | /supɚ/ | *super* | /suna/ | 'sand' |
| | | | /supuN/ | 'spoon'[2] |
| ʃi | /ʃild/ | *shield* | /ʃika/ | 'deer' |
| | /ʃɪp/ | *ship* | /ʃippo/ | 'tail'[1] |
| | /ʃip/ | *sheep* | /ʃiso/ | 'seesaw' |
| ʃe or ɕe | /ʃɛl/ | *shell* | | |
| | /ʃevɪŋ/ | *shaving* | | |
| | /ʃep/ | *shape* | | |
| ʃa or ɕa | /ʃɑp/ | *shop* | /ɕamodʑi/ | 'rice paddle' |
| | /ʃɑt/ | *shot* | /ɕatsu/ | 'shirt' |
| | /ʃɑɹp/ | *sharp* | /ɕawa/ | 'shower' |
| | | | /ɕaɕiŋ/ | 'photograph'[2] |
| | | | /ɕaberu/ | 'chat, talk'[2] |
| ʃo or ɕo | /ʃoldɚ/ | *shoulder* | /ɕo:dʑi/ | 'paper screen' |
| | /ʃo/ | *show* | /ɕokupaN/ | 'bread' |
| | /ʃoɹp/ | *short* | /ɕo:ju/ | 'soy sauce' |
| ʃu or ɕu | /ʃu/ | *shoe* | /ɕu:mai/ | 'Chinese dumpling' |
| | /ʃut/ | *shoot*[1] | /ɕukurimu/ | 'creme puff'[1] |
| | /ʃʊgɚ/ | *sugar* | /ɕudzu/ | 'shoes'[1] |

*Note*: A small number of stimuli were used only with children[1] or only with adults[2].

# References

Akamatsu, T. (1997). *Japanese phonetics: Theory and practice*. Newcastle: Lincom Europa.

Baum, S. R., & McNutt, J. C. (1990). An acoustic analysis of frontal misarticulation of /s/ in children. *Journal of Phonetics, 18*, 51–63.

Behrens, S. J., & Blumstein, S. E. (1988). Acoustic characteristics of English voiceless fricatives: A descriptive analysis. *Journal of Phonetics, 16*, 295–298.

Boersma, Paul & Weenink, David (2005). Praat: Doing phonetics by computer (version 5.0.24) [Computer program]. Retrieved April 17, 2005, from ⟨http://www.praat.org/⟩.

Dinnsen, D. A. (1992). Variation in developing and fully developed phonetic inventories. In C. Ferguson, L. Menn, & C. Stoel-Gammon (Eds.), *Phonological development: Models, research, implications* (pp. 191–210). Timonium, MD: York Press.

Fernandez, S., Feijoo, S., Balsa, R., & Barros, N. (2000). Perceptual effects of coarticulation in fricatives. *In Proceedings of the IEEE international conference on acoustics, speech, and signal processing, 3*, 1347–1350.

Fletcher, S., & Newman, D. (1991). [s] and [ʃ] as a function of linguapalatal contact place and sibilant groove width. *Journal of the Acoustical Society of America, 89*, 850–858.

Forrest, K., Weismer, G., Hodge, M., Dinnsen, D. A., & Elbert, M. (1990). Statistical analysis of word-initial /k/ and /t/ produced by normal and phonologically disordered children. *Clinical Linguistics and Phonetics, 4*, 327–340.

Forrest, K., Weismer, G., Milenkovic, P., & Dougall, R. N. (1988). Statistical analysis of word-initial voiceless obstruents: Preliminary data. *Journal of the Acoustical Society of America, 84*, 115–123.

Funatsu, S. (1995). Cross language study of perception of dental fricatives in Japanese and Russian. *Proceedings of the International Congress of Phonetic Sciences, 4*, 124–127.

Gafos, A. (1999). *The articulatory basis of locality in phonology*. Outstanding Dissertations in Linguistics. New York: Garland.

Halle, M., & Stevens, K. N. (1997). The postalveolar fricatives of Polish. In H. H. a. H. F. Shigeru Kiritani (Ed.), *Speech production and language: In honor of Osamu Fujimura*, Vol. 13 (pp. 176–191). Berlin, New York: Mouton de Gruyter.

Hughes, G. W., & Halle, M. (1956). Spectral properties of fricative consonants. *Journal of the Acoustical Society of America, 28*, 303–310.

Ingram, D. (1989). *First language acquisition: Method, description and explanation*. Cambridge: Cambridge University Press.

Jakobson, R. (1941/1968). *Child language, aphasia, and phonological universals*. Mouton: The Hague.

Jongman, A., Wayland, R., & Wong, S. (2000). Acoustic characteristics of English fricatives. *Journal of the Acoustical Society of America, 108*, 1252–1263.

Kent, R. (1992). The biology of phonological development. In C. Ferguson, L. Menn, & C. Stoel-Gammon (Eds.), *Phonological development: Models, research, implications* (pp. 65–90). Timonium, MD: York Press.

LaRiviere, C. (1975). The distribution of perceptual cues in English prevocalic fricatives. *Journal of Speech and Hearing Research, 18*, 613–622.

Li, F. (2008). *The phonetic development of voiceless sibilant fricatives in children speaking English, Japanese and Mandarin Chinese*. Doctoral Dissertation, Ohio State University.

Li, F., & Edwards, J. (2006). *Contrast and covert contrast in the acquisition of /s/ and /ʃ/ in English and Japanese*. Poster presented at the 10th laboratory phonology conference, June 29–July 1, Paris, France.

Locke, J. L. (1983). *Phonological acquisition and change*. New York: Academic Press.

Macken, M. A., & Barton, D. (1980). A longitudinal study of the acquisition of the voicing contrast in American-English word-initial stops, as measured by voice onset time. *Journal of Child Language, 7*, 41–74.

Maddieson, I., & Precoda, K. (1990). *UPSID-PC. The UCLA Phonological Segment Inventory Database*. [MS-DOS package available online from the UCLA Phonetics Laboratory at ⟨http://www.linguistics.ucla.edu/faciliti/sales/software.htm⟩].

Maxwell, E. M., & Weismer, G. (1982). The contribution of phonological, acoustic, and perceptual techniques to the characterization of a misarticulating child's voice contrast for stops. *Applied Psycholinguistics, 3*, 29–43.

Munson, B., Li, F., Yoneyama, K., Hall, K., Beckman, M.E., Edwards, J., & Sunawatari, Y. (2008). *Sibilant fricatives in Japanese and English: different in production or perception?* Paper presented at the Annual Meeting of the Linguistics Society of America, January 4–6, Chicago, IL.

Nakanishi, Y., Owada, K., & Fujita, N. (1972). Koon kensa to sono kekka no kosatsu [Results and interpretation of articulation tests for children]. RIEEC Report [Annual Report of Research Institute Education of Exceptional Children, Tokyo Gakugei University], Vol. 1 (pp. 1–41).

Nissen, S. L., & Fox, R. A. (2005). Acoustic and spectral characteristics of young children's fricative productions: A developmental perspective. *Journal of the Acoustical Society of America, 118*, 2570–2578.

Nittrouer, S. (1995). Children learn separate aspects of speech production at different rates: Evidence from the spectral moments. *Journal of the Acoustical Society of America, 97*, 520–530.

Nittrouer, S., Studdert-Kennedy, M., & McGowan, R. (1989). The emergence of phonetic segments: Evidence from the spectral structure of fricative-vowel syllables spoken by children and adults. *Journal of Speech and Hearing Research, 32*, 120–132.

Prather, E., Hedrick, D., & Kern, D. (1975). Articulation development in children aged two to four years. *Journal of Speech and Hearing Disorders, 403*, 179–191.

Raudenbush, S. W., Bryk, A. S., Cheong, Y. F., & Congdon, R. (2004). *HLM 6: Hierarchical linear and nonlinear modeling*. Lincolnwood, IL: Scientific Software International.

Scobbie, J. E., Gibbon, F., Hardcastle, W. J., & Fletcher, P. (2000). Covert contrast as a stage in the acquisition of phonetics and phonology. In M. Broe, & J. Pierrehumbert (Eds.), *Papers in laboratory phonology V: Language acquisition and the lexicon* (pp. 194–203). Cambridge: Cambridge University Press.

Scobbie, J. M. (1998). Interactions between the acquisition of phonetics and phonology. In M. C. Gruber, D. Higgins, K. Olson, & T. Wysocki (Eds.), *Papers from the 34th annual regional meeting of the Chicago Linguistic Society*, Vol. II (pp. 343–358). Chicago: Chicago Linguistics Society.

Shadle, C.H. & Mair, S.J. (1996). *Quantifying spectral characteristics of fricatives*. Paper presented at the the international conference on spoken language processing, Philadelphia, PA.

Smit, A. B., Hand, L., Freilinger, J. J., Bernthal, J. E., & Bird, A. (1990). The Iowa articulation norms project and its Nebraska replication. *Journal of Speech and Hearing Disorders, 55*, 779–798.

Stevens, K. N. (1998). *Acoustic phonetics*. Cambridge, MA: MIT Press.

Stone, M., Faber, A., Raphael, L., & Shawker, T. (1992). Cross-sectional shape and linguopalatal contact patterns in [s], [S], and [l]. *Journal of Phonetics, 20*, 253–270.

Tabain, M. (2001). Variability in fricative production and spectra: Implications for the hyper- and hypo- and quantal theories of speech production. *Language and Speech, 44*(Part 1), 57–94.

Templin, M. (1957). *Certain language skills in children*. Minneapolis: University of Minnesota Press.

Toda, M., & Honda, K. (2003). *An MRI-based cross-linguistic study of sibilant fricatives*. Paper presented at the sixth international seminar on speech production, December 6–10, Manly, Australia.

Tsurutani, C. (2004). Acquisition of Yo-on (Japanese contracted sounds) in L1 and L2 phonology in Japanese second language acquisition. *Journal of Second Language, 3*, 27–47.

Tyler, A. A., Figurski, G. R., & Langdale, T. (1993). Relationships between acoustically determined knowledge of stop place and voicing contrasts and phonological treatment progress. *Journal of Speech and Hearing Research, 36*, 746–759.

Weismer, G., & Elbert, M. (1982). Temporal characteristics of functionally misarticulated /s/ in 4- to 6-year-old children. *Journal of Speech and Hearing Research, 25*, 275–287.

Whalen, D. H. (1984). Subcategorical phonetic mismatches slow phonetic judgments. *Perception and Psychophysics, 35*, 49–64.

White, D. (2001). Covert contrast, merger, and substitution in children's productions of /k/ and /t/. Unpublished Master's Thesis, The Ohio State University.

Yasuda, A. (1970). Articulatory skills in three-year-old children. *Studia Phonologica, 5*, 52–71.