

# Formant-cavity affiliation in sibilant fricatives

*Martine Toda, Shinji Maeda and Kiyoshi Honda*

## Introduction

Sibilant fricatives form a subgroup of fricatives that outputs high intensity noise. This high intensity reflects the mechanism of noise generation where the airflow is directed against an obstacle such as incisors (Shadle 1985, 1991). This mechanism is also supported by the high jaw position of sibilants (Lee, Beckman and Jackson 1994; Mooshammer, Hoole and Geumann 2007), implying that the position of lower incisors with respect to the airstream is relevant for the noise generation. The intense frication noise is known to carry perceptual cues for the place of articulation of sibilants (e.g. English /s/ vs. /ʃ/, Harris 1958). Contrary to stops, nasals and non-sibilant fricatives, sibilants are thus the rare consonants whose place as well as manner cues are signaled primarily by the spectral structure of the segment itself.

The acoustic characteristics of fricatives differ however from those of vowels in many aspects. Firstly, the spectral envelope of the noise source can maintain a high intensity level, up to about 15 kHz for some utterances, whereas the voiced source of vowels damps out rapidly above 5 kHz. Secondly, since the generation of fricatives' source requires a tight constriction, the back cavity located behind the constriction tends to be acoustically inactive. Therefore, the formant structure of sibilant fricatives is primarily determined by the front cavity resonances. Thirdly, fricatives generate both resonances and antiresonances, because of the noise source to be located midway in the vocal tract, in particular within the front cavity, near the teeth (Shadle 1991 for [s] and [ʃ]). The antiresonances, like resonances, contribute in shaping the overall spectrum. To summarize, while a source-filter independence can be assumed in obstacle-type fricatives (Shadle 1985: 178), sibilants' spectrum differs from that of vowels not only by the nature of excitation sources (noise source opposed to the periodic glottal source) but also by its specific spectral structure.

Fricatives' noise spectra are often described by global acoustic parameters such as the center of gravity, peak frequency, spectral tilt, etc.

However, such descriptions are insufficient when the spectral structure is to be interpreted in articulatory terms.

This chapter attempts an analytical description of the sibilants' spectral structure in order to clarify the articulatory-to-acoustic mapping in sibilant fricatives. After a short summary of the literature about the noise source, the general acoustic (resonance) properties of simple fricative-like models will be discussed and illustrated through acoustic simulations. Then, three-dimensional vocal tract shape of Polish sibilants will be examined in teeth-inserted high-resolution magnetic resonance imaging (MRI) data, in order to provide a detailed description of a variety of sibilant fricatives. Finally, we will attempt to clarify the formant-cavity affiliation in Polish /ç/ and /ʂ/ by running acoustic simulations with realistic and modified models derived from the MRI data. The results are discussed in relation with the natural spectra.

## 1. Fricatives' noise source

Several types of noise sources are assumed to exist in fricatives. When an air jet forms at the constriction and flows into a wider cavity, such as the front cavity, the laminar flow turns into turbulences due to the viscosity of the air, at a certain threshold indicated by Reynolds' number (see Shadle's chapter, this volume). Those turbulences give rise to monopole and quadrupole noise sources (Pastel 1987:21). The effect of quadrupole sources is however considered to be minor in the speech signal.

In sibilant fricatives [s] and [ʃ], in addition to the sources resulting from a free jet just described, *obstacle* sources arise at the obstacle where the air jet exiting from the constriction impinges. The upper or lower incisors, depending on the model (Shadle 1991), constitute the obstacle. An obstacle configuration gives rise to dipole sources. The noise intensity is amplified and spreads over higher frequencies when an obstacle source is involved (Stevens, 1998: 107).

Dipole sources can be understood as two monopole sources pulsating in opposite phase, one releasing, the other absorbing [air] mass (Pastel 1987: 21). Those sources are oriented perpendicular to the obstacle (Curle 1955). It has been shown that dipole sources couple the best to the principal acoustic mode along the length of the vocal tract, when the obstacle is oriented perpendicular to it (Pastel, 1987).

In [ç] and [x], Shadle (1985:178) has showed that the noise source could well be approximated by dipole sources distributed along a surface

downstream from the constriction. In these articulatory configurations, the air jet exiting from the constriction impinges the palatal roof, but it is not perpendicularly oriented with respect to the obstacle. Therefore, the resulting noise level might be lower than in case the obstacle is oriented perpendicularly (Stevens 1998:102). Nevertheless, it can be suggested that such a source generation mechanism plays an important role in the fricatives that are characterized by a long constriction formed by a domed tongue in general. Fant (1970) notes that the best fit between acoustic modeling and actual sound spectrum in ‘š’ ( [ʃ] ) is obtained when the source is placed within the palato-alveolar constriction. Narayanan and Alwan (2000) also obtained an optimal fit of the simulated spectrum to the actual noise spectrum of English /ʃ and ʒ/ by using different source types as (1) a *wall* dipole source in the vicinity of constriction, in addition to (2) another dipole source located at the teeth, and (3) a constant monopole source located at constriction exit, as well as (4) a voiced source for /ʒ/. The authors note that these dipole sources greatly influenced the spectral shape over most of the frequency range, i.e. up to 10 kHz. Therefore, it can be conceived that sibilants with a similarly long constriction also involve, at least, the wall-obstacle source (near the constriction) in addition to the teeth-obstacle source.

Dipole source is considered to have relatively flat long-term spectrum with a single broad peak (see Stevens 1998: 103), and its central frequency as well as the left and right slopes seemingly vary according to the subject or to the identity of the fricative (Narayanan and Alwan’s 2000 three parameter dipole source models, see Figure 1). Although the fricatives’ spectral shape is the product of the source envelope and the vocal tract transfer function, the latter largely determines the spectral characteristics of the sibilant fricatives.

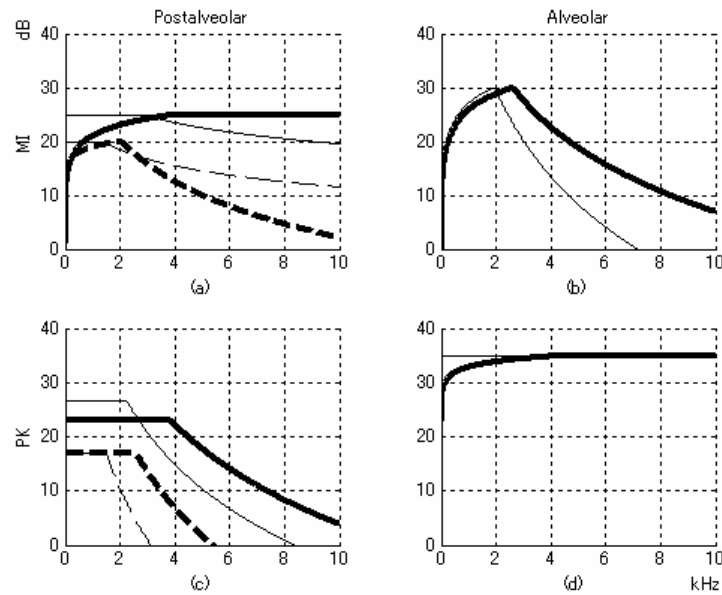


Figure 1. Graphical representation of three parameter dipole source models formulated by Narayanan and Alwan's (2000) for English fricatives. The upper graphs are for subject MI (a and b); the bottom graphs for subject PK (c and d). Left ones are for postalveolar /ʃ/ and /ʒ/ (a and c), and right ones for alveolar /s/ and /z/ (b and d). In each graph, the curves corresponding to the voiceless fricatives are drawn in thick line and those for the voiced fricatives in thin line. Plain lines correspond to the teeth dipole source, broken lines to the wall dipole source. The monopole and voiced sources are not represented.

## 2. Theoretical considerations: Acoustic structure of fricatives

Consonants possess an underlying formant pattern, i.e., F-pattern defined by Fant (1970: 25) as do vowels, which explains the formant transitions at consonant-vowel boundaries. However, the F-pattern is not entirely visible in the acoustic realization of the consonants. In fricatives, the front cavity resonances excited by a supraglottal source are thought to determine the overall spectral shape, because of the existence of a tight constriction. In

addition, the transfer function of fricatives is characterized by antiformants of various origins as will be developed below. In passing, we note in this paper that only the visible formants will be referred to as fricative formants F1, F2, etc., regardless of their relation to the adjacent vowel formants.

## 2.1. Method

The simplest fricative model, similar to Flanagan (1972:74), is a straight-tubed tract that contains a constriction. In an [s]-like configuration, the narrow tube, which mimics the tongue constriction, the small front cavity and the lip opening altogether, would be put at the frontmost portion of the model, as shown in Figure 2.

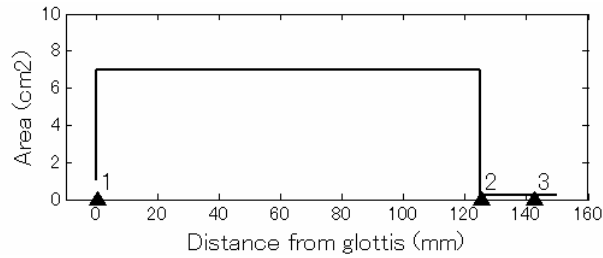


Figure 2. A simple [s]-like fricative model. Numbered black triangles indicate alternative source locations. Back cavity length = 12.5 cm; back cavity cross-sectional area = 7 cm<sup>2</sup>; front cavity length = 2.5 cm; front cavity cross-sectional area = 0.2 cm<sup>2</sup>.

It is possible to calculate the transfer function of such a vocal tract model by using acoustic simulation. The program VTF\_fric, a frequency domain version of VTCales (Maeda 1982), is used in this study. This program takes an area function as input, as schematized in Figure 2, and the parameters that are listed in Table 1 for default values.

Table 1. Default parameter values used with the acoustic simulation program VTF\_fric.

<i>wall properties</i>	<i>radiation load (at lips)</i>	<i>subglottal system</i>	<i>source type</i>
yielding	RL circuit	OFF	pressure
<i>output type</i>	<i>glottal area</i>	<i>length of sections</i>	<i>frequency samples</i>
radiated pressure	1 cm <sup>2</sup>	1 mm	every 10 Hz

The frequency range of our simulations comprises up to 24 kHz. This unusually wide frequency range permits the visualization of the higher resonances of the fricative models and thus helps us to correctly interpret the resonance modes. However, the higher frequencies (8 kHz and up) of the simulated transfer functions should not necessarily match the actual noise spectra, since in a three-dimensional real vocal tract, cross modes or higher order modes (i.e., resonance modes that are not longitudinal with respect to the vocal tract) can also occur at such high frequencies, which is naturally not predicted by this simulation method that assumes a plane-wave propagation.

## 2.2. Natural resonance frequencies of fricative models

When the two-tube resonating system is excited by a source located at the glottis (e.g., source location 1 in Figure 2), the transfer function (Figure 3, upper curve) is characterized only by poles, that correspond to the natural frequencies of the model.

This all-pole function involves three series of resonances. The lowest resonance of this system is a Helmholtz resonance, expected around 170 Hz. This resonance is not noticeable on the transfer function because of the important damping of the low frequencies due to the glottal opening. The resonances appearing with the shortest interval are those of the 12.5 cm long back cavity. Since the both ends of the back cavity are almost closed, we expect  $n/2$  wavelength resonances around 1.4, 2.8, 4.2 kHz and so on, which roughly fits the observed interval of 1.3 kHz. The resonances of the front cavity of 2.5 cm long are superimposed to the back cavity resonances. Since the both ends of this cavity are roughly open, we expect multiples of half wavelength resonances around 7, 14, 21 kHz, and so on. Actually, the first resonance is observed around 6.5 kHz. This value is slightly lower than expected because of the effect of lip radiation. The second and third resonances overlap to the back cavity resonances around 13 and 19.5 kHz, respectively, making their relative intensity greater.

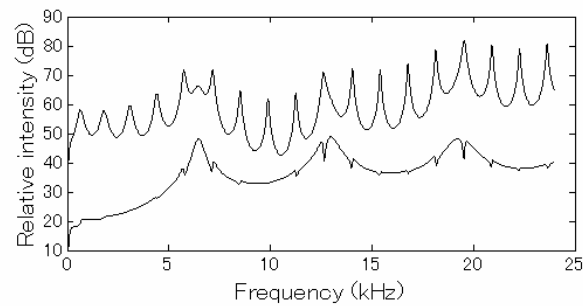


Figure 3. Transfer functions of a simple /s/ model. The upper curve (displayed with +30 dB offset for clarity) corresponds to the source location 1 (glottis), as shown in Figure 2. The lower curve corresponds to the source location 2 (at the junction of back and front cavities).

### 2.3. Source located in front of a vocal tract cavity

When an excitation source is located downstream of a vocal tract cavity, this cavity produces pole-zero pairs in the transfer function. For example, when the noise source is located within or at the entrance of the front cavity as indicated by the triangle 2 in Figure 2, the back cavity produces pole-zero pairs. The paired poles and zeros annihilate each other in a theoretical condition where the interaction between the front and back cavities is minimal, that is, when the difference of area between these cavities is large. In real conditions, the vocal tract cavities are not completely independent, and the pole-zero pairs will not annihilate each other completely. The tighter the constriction the closer will be the distance between the paired pole and zero, so that their influence on the overall spectrum will tend to be negligible as seen in Figure 3, bottom curve. For straight tube models, an area ratio of 1 to 10 is usually considered to be sufficiently large that the interaction between two adjacent cavities can be ignored.

### 2.4. Source located midway within a cavity

When an excitation source is located midway within a cavity, this cavity produces free poles and free zeros (Figure 4). This applies to the front cavity when the source is located, actually, at 8 mm from the constriction exit (source 3 of Figure 2). The zeros occur due to multiple reflections of sound waves in the cavity at the frequencies where the phase reversal takes

place between forward and backward waves to cancel each other. Because of the inversion of phase in the forward/backward direction, dipole pressure sources (correctly oriented with respect to the resonance mode) do not cause the same response as monopole flow sources (Figure 4). Notice that the frequency of free poles remains the same independently of the source types, as well as the source locations (cf. Figure 3). Also, the frequency interval of the free zeros is not related to the frequency of poles, but to the position of the excitation source within the front cavity.

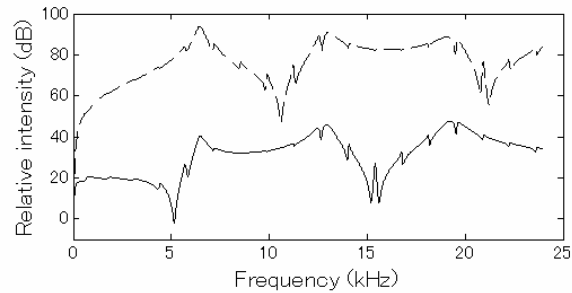


Figure 4. Transfer function of a simple /s/ model excited by a source located midway within the front cavity (8 mm behind the lip opening, numbered 3 in Figure 2). Top: Flow source equivalent to monopole. Bottom: Pressure source equivalent to dipole. The relative intensity is defined with respect to the first frequency sample (10 Hz) set to 0 dB.

If the noise source were distributed along a cavity, as is the case of a long and tight constriction in [x], the frequency of free zeros would also be distributed along the frequency axis. If this is the case, the influence of free zeros on the overall spectrum would be insignificant in comparison with that of free poles, whose frequency remains constant regardless of the location of the sources. In [s] and [ʃ], since the source is likely to be localized around the incisors (Shadle 1985; 1991), it is more likely that the noise spectrum would be affected by sharp free zeros.

## 2.5. Vocal tract side branches

The location of the source is not the only factor that can introduce zeros in the spectrum of fricative-like configurations. In some articulatory settings, typically in [ʃ or ʂ], the front cavity is enlarged by a sublingual space created by a raised tongue tip combined with a retracted position of the



tongue. In such a case, especially when the sublingual cavity is large, the lowest resonance arising from the front cavity will be best modeled by considering the sublingual space as a side branch. A side branch in an acoustic system creates the condition for a sound wave to bifurcate, and reunify at the observation point with phase reversal, where zeros appear in the transfer function in the same manner with the case of a source located midway within a vocal tract cavity.

The configuration schematized in Figure 5 mimics a case for a post-alveolar fricative, where the main tract comprises three tubes: the back cavity, the constriction and the front cavity. Its total length is now 17 cm, the constriction being of 2.5 cm long (same as the previous model, Figure 2) and the front cavity measuring 2 cm. A side branch of 1.5 cm is connected at the posterior end of the front cavity. The source is located 1 cm behind the lip exit, as indicated by the triangle.

Theoretically, a side branch introduces additional poles and zeros in the transfer function. However, since the area of the sublingual cavity is likely to be about the same as that of the front cavity at the branching point, a strong interaction is expected between the side branch and the front cavity. The sublingual cavity is therefore likely to modify the pole and zero frequencies of the front cavity as seen in this simulation (Figure 5, bottom), where the lowest peak shifts from about 3 kHz to 2 kHz between the not-branching (broken line) and branching (plain line) configurations. It is worth noting that the effect of the sublingual cavity is very similar to that of lengthening the front cavity. The dotted line represents the transfer function of a straight model, without a side branch, where the front cavity measures 3.5 cm (the total length and constriction length are kept the same). The main difference is found in the frequency of the back cavity pole-zero pairs.

Another kind of side branches consists of the interdental space. These lateral cavities are morphologically delimited by tongue sides, upper and lower teeth, and cheeks. These side branches are connected to the main vocal tract at the lateral sides of the front cavity. Although their volume, and thus the magnitude of their resonances are assumed to be rather small, they may have some impact on the spectrum during the vowel-fricative transitions, creating formant discontinuities due to the sudden coupling/decoupling of their poles and zeros.

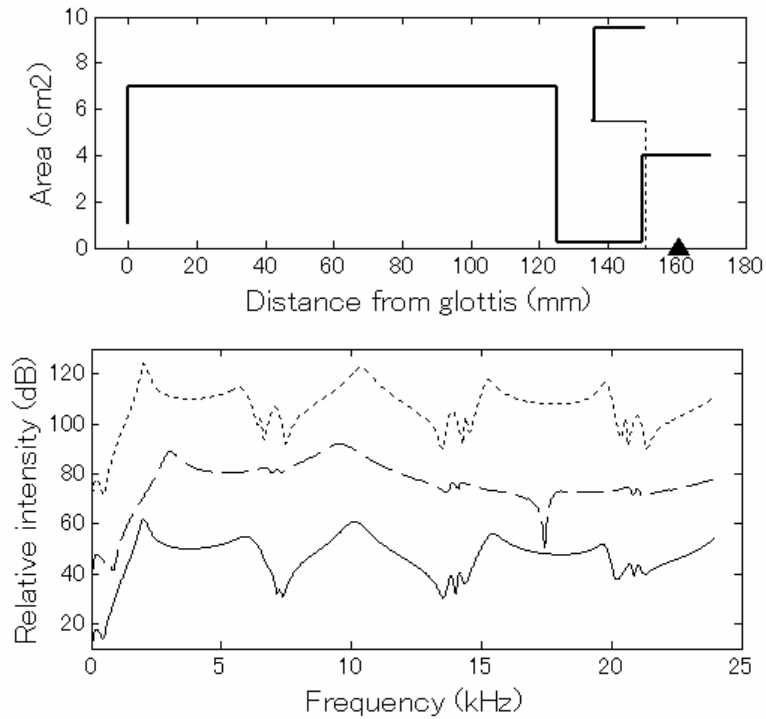


Figure 5. Top: area function of a branching post-alveolar fricative model. The side branch (15 mm long) is connected to the main tract at the posterior end of the front cavity. Bottom: its transfer function. The transfer function of the model without a side branch is given in broken line. That of a straight (non-branching) model which front cavity measures 35 mm is given in dotted line. All of the transfer functions are obtained with a pressure source located 10 mm behind the lip exit. The curves are shifted by 30 dB offsets for a better visibility.

## 2.6. Generalization

To summarize, the fricatives' transfer function is likely to be made up of the following constituents:

- Pole-zero pairs of independent vocal tract cavities (e.g. back cavity) excited by a source located downstream (within the front cavity), having a minor influence on the overall spectral shape;

- An appreciable shift of front cavity poles and zeros due to the connection of a sublingual cavity side branch, in postalveolars;
- Free zeros and free poles arising from cavities with a noise source located within it. In typical [s] or [ʃ/ʂ] configurations, the spectra of sibilant fricatives are likely to be characterized by a sharp cut-off at the left side of the first fricative formant because of a free zero of this kind. Other sharp dips are also likely to occur in frequencies above the main spectral peak or peaks;
- Free poles, not accompanied by zeros, arise from cavities excited by a source located upstream. In an [ʃ/ʂ]-like model, consisting of three tubes instead of two (wide back cavity tube, narrow constriction tube, and wide front cavity tube), sources located within the constriction tube will make the front cavity to produce free poles. Some [s] can also be represented by the three tube model, with different lengths. The third tube, wider than the second constriction tube, represents now the lip cavity. With a noise source located at the incisors (at the end of the constriction tube), the lip tube, despite its short length, will be responsible for free poles of high, but audible frequencies.

### **3. Vocal tract shapes of sibilant fricatives**

Recently, acoustic study on fricatives have benefited from the advancement of the medical imaging techniques, especially the magnetic resonance imaging (MRI). In earlier times, the vocal tract area function of fricatives was difficult to estimate from the sagittal profile derived from x-ray data, and needed to be completed with separate data such as x-ray tomography (e.g. Shadle, 1991). Moreover, the derived area functions were optimized so that the calculated vocal-tract transfers better match with the measured spectra of the corresponding fricatives (Fant, 1970). In the investigation of the acoustic system of fricatives, data of particularly high accuracy are needed, because of the presence of constrictions with a small dimension. This is because small errors in the small structures can exaggeratedly affect the estimated formant frequencies. In Narayanan, Alwan and Haker (1995), three-dimensional data of English fricatives were published for the first time, including, in particular, the shape as well as the volume of the sublingual cavity in /ʃ/. However, the accuracy of the extracted vocal tract contour in the front cavity region needed improvement, especially for acoustic study purposes.

More recently, we acquired a large amount of voiceless sibilant fricative MRI data (30 subjects, Toda forthcoming), from which we extracted a few samples here. In this study, the priority was given to the accuracy of the data in the front cavity region, comprising the complex teeth contour as well as the narrow tongue constriction.

In order to correctly apprehend the vocal tract shape of sibilant fricatives, the teeth contour could by no means be ignored. It used to be, however, a challenge to combine the teeth contour to the MR images, and several approaches have been experimented, either combining hardware teeth casts to their MRI data, or calling for a coating medium (Narayanan, Alwan and Haker, 1995), or using post-hoc image processing (Takemoto *et al.* 2004). In the present study, a manual post-hoc image processing procedure was adopted. The main advantage of this method is the high precision with which the teeth contour can be combined to the articulatory data.

The Polish sibilants will be analyzed in the present study. The Polish language possesses three voiceless sibilants with contrastive ‘places of articulation’, which constitute an interesting sample of sibilant sounds that the human vocal tract is able to produce in robust ways. Especially, Polish possesses two kinds of postalveolar sibilant fricatives, /ʃ/ and /ç/, radically different in their tongue shape, /ç/ being palatalized, as reported in Halle and Stevens (1997) based on Wierzchowska’s (1965) data. From the acoustic point of view, the frication noise’s centers of gravity of the two sounds overlap with each other (Jassem 1995: 3 male subjects; Nowak 2006: 1 female subject ; 1 male subject in Zygis and Hamann 2003), or they are lower for /ʃ/ (1 female subject in Zygis and Hamann 2003); for both fricatives, the frication noise seems not far different from that of English /ʃ/. Nevertheless, the frication noise appears to be perceptually distinctive when these fricatives are presented in isolation (Lisker 2001 and Nowak 2006). The F2 transitions, as shown in Figure 6, are very different between /ʃ/ and /ç/, where the latter has a higher locus. The transitional information, especially of the following vowel, has a strong effect on native listners’ judgement about the identity of the sibilant (Nowak 2006).

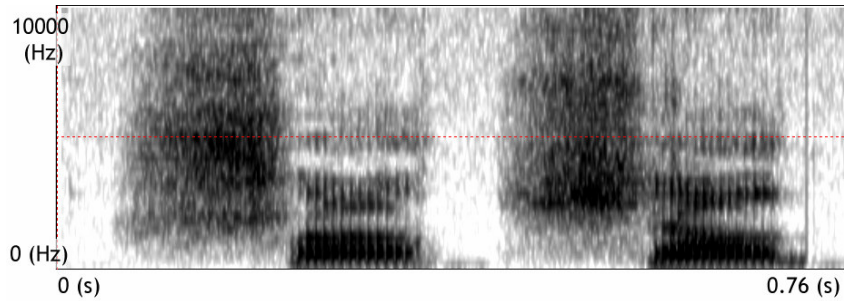


Figure 6 Spectrogram of the initial syllables [ʂa] and [ʧa] extracted, respectively, from the words *szatan* and *siadam*, uttered by subject P2.

According to the UPSID database (Maddieson and Precoda 1989), tongue shape contrast (involving palatalization, dental/alveolar or palatal/retroflex/ palato-alveolar contrasts) is one of the most common ways of contrasting sibilant fricatives in the world's languages, following voiced/voiceless and anterior/non anterior contrasts. For instance, in the languages having an inventory of two or more sibilant fricatives (304 languages out of 451), 191 involve a contrast concerning 'place' (in the general sense covering tongue position as well as tongue shape), whereas, for comparison, 147 involve a voiced/voiceless contrast. Among the 191 languages involving a place-related feature, 180 languages possess an anterior/non anterior contrast, and 39 a tongue shape contrast. In languages having more than 3 sibilants, anterior/non anterior and tongue shape contrasts can be combined.

Therefore, it is of interest to examine whether the variation in tongue shape results in specific characteristics of the frication noise spectrum in addition of that of formant transitions. In particular, Halle and Stevens (1997) suggested that the presence of a long and thin palatal channel should reduce the acoustic coupling of the back cavity in /ʧ/, in opposition to /ʂ/. Also, according to the same authors, the palatal channel should provide its own formants and antiformants, which also contribute to the spectral difference between /ʧ/ and /ʂ/.

### 3.1. Method

Two native Polish male subjects participated in the experiment. The data acquisition required two sessions: (1) acoustic recordings and (2) MRI

acquisition. The acoustic recordings are described in the following section (5.1.1.).

### 3.1.1. MRI acquisition

The MRI data were acquired at the Brain Activity Imaging Center in ATR (Kyoto, Japan), with a Shimadzu-Marconi Eclipse 1.5T scanner. The MRI parameters of the acquisition are given in Table 2. The subjects lied down in supine position in the scanner. In this study, in order to obtain the most *natural* data as possible, special efforts have been made to keep the scan duration as brief as possible (23 seconds). A high sensitivity custom-made coil (Takano *et al.* 2003) was used to obtain high resolution data of the orofacial region with a good signal to noise ratio despite the short acquisition time.

Table 2. MRI acquisition parameters

<i>scan direction</i>	<i>field of view</i>	<i>resolution (1 pixel =)</i>	<i>slice thickness and spacing</i>
sagittal	128 mm × 128 mm × 45 mm	0.25 mm × 0.25 mm	1.5 mm
<i>TE</i>	<i>TR</i>	<i>scan duration</i>	<i>sequence name</i>
3.3 ms	10 ms	23 s	Fast 3D

The subjects were presented words or word sequences containing the target fricatives in various phonetic contexts (2 or 3 vowel contexts for each fricative), and were told to produce a sustained fricative *as if* it was realized in those contexts. They were instructed to keep producing the fricative, during the whole duration of the scan, and maintain the same posture when they could not sustain it to the end. These instructions are to prevent the blurring of the MR images due to movement. The subjects were already accustomed to the corpus, and trained for the sustained phonation during the acoustic recording session. They rehearsed again immediately before the MRI session took place. The corpus was printed in large characters and glued inside the scanner, so that the subjects could refer to it during the whole MRI session.

The consistency of the articulation within the sibilant phonemes was verified through the comparison of the raw MRI data corresponding to the various phonetic contexts. The sustained fricatives /s/, /ç/ and /ʃ/ presented in the words *sadza* [sadza], *siadam* [çadam] and *szatan* [ʃatan] will be analyzed in this study.

In addition, special *teeth scans* have been acquired, in order to extract the subjects' numerical teeth casts (usually, the teeth appears in the same brightness as the air and thus their contour is not visible in MR images). For that scan, the subjects were instructed to put their tongue and lips tightly close to their teeth. The same acquisition parameters as the fricative scans were used in order to insure the uniformity of the data and facilitate the post-processing.

### *3.1.2. Teeth cast extraction and insertion*

The numerical teeth casts were extracted from the MRI teeth scans through brightness inversion and by manually selecting the regions of interest, including bones (mandible and hard palate).

Then, the upper and lower teeth casts were merged manually into the fricative scans by using a medical imaging software (Intage-rv), through rotations and translations. The details of bones were of great help in the insertion of these dental casts into each of the sibilant dataset with good accuracy. The precision of this procedure is estimated to be around 0.5 mm and 1 degree.

### *3.1.3. Measurement of vocal tract area functions*

The vocal tract area function of sibilants was measured. Considering the vocal tract to be roughly horizontal in the front oral cavity region, the area of the airway was measured on coronal images (perpendicular to the original sagittal images) at an interval of 1 mm. This method lacks accuracy in the two measures: (a) the length of the lip tube, where the side boundaries of the airway at its extremity are not well captured because of its curvature (For the subject P1, no lip cavity could be measured at all); and (b) the length of the front cavity comprising the sublingual portion and the length of the tongue constriction because of their angle with respect to the coronal slices. Both of these measurement errors would lead to shorter lengths than the actual vocal tract, and thus a shift of the resonances towards higher frequencies in the estimated transfer functions. However, we assume that the overall resonance pattern will not be significantly affected by these inaccuracies, since the structural relationships among the vocal tract cavities are preserved.

## 3.2. Results

The mid-sagittal images as well as coronal images showing the constriction and the tongue dorsum are given in Figures 7-8.

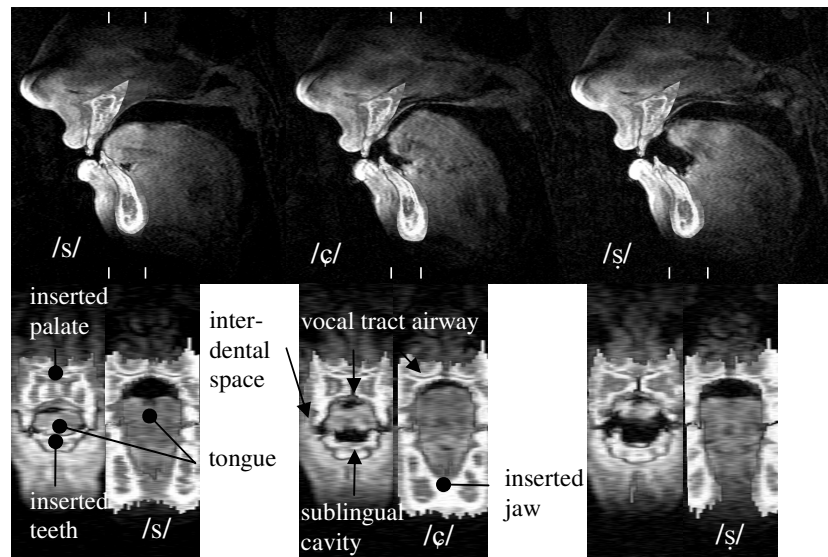
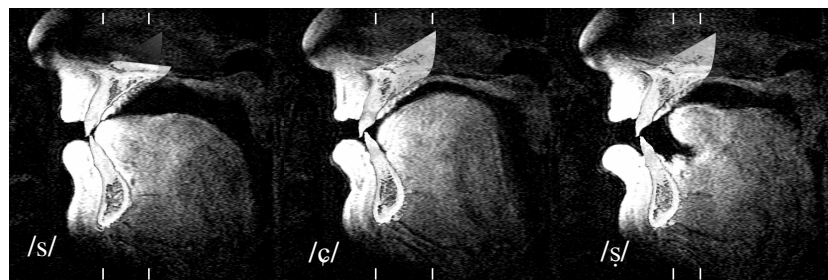


Figure 7. Mid-sagittal (top) and coronal (bottom) views of the Polish sibilants /s/ (left), /ç/ (center) and /ʂ/ (right), uttered by P1. Sustained fricatives produced in an imaginary [ɑ] context. The position of coronal slices are indicated by tick marks in the corresponding sagittal images.





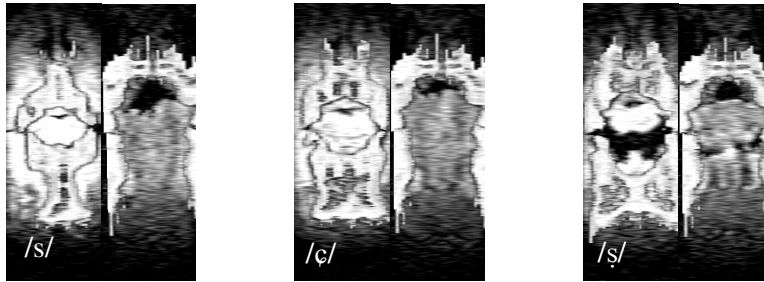


Figure 8. Mid-sagittal (top) and coronal (bottom) views of the Polish sibilants /s/ (left), /ç/ (center) and /ʂ/ (right), uttered by P2. Sustained fricatives produced in an imaginary [a] context. The position of coronal slices are indicated by tick marks in the corresponding sagittal images.

From Figures 7 and 8, it appears that subject P1 and P2 produced the three Polish sibilants with certain common characteristics.

/s/ has a very front constriction, where the tongue is in contact with the internal surface of the lower incisors. The tongue is the most retracted in /ʂ/, with the narrowest pharyngeal cavity, especially for P2. A large sublingual cavity is noticeable for /ʂ/ both on the sagittal as well as coronal images. Although P1 exhibits a sublingual cavity also for /ç/, its volume is smaller than in /ʂ/. It can also be remarked that in /ʂ/, the depth of the sublingual cavity is superior to the distance between the tightest point of the tongue constriction to the teeth. The lowest front cavity resonance (other than Helmholtz resonance) should therefore involve the vertical dimension of the front cavity, including the sublingual portion. P2 realizes a gradual lip protrusion in the order /s/ < /ç/ < /ʂ/, whereas no particular lip gesture is noticeable for P1. The increasing front cavity size /s/ < /ç/ < /ʂ/ can also be seen on the area functions given for P2 (Figure 9).

/ç/ is characterized by a strong tongue dorsum doming in both subjects. This sagittal doming is accompanied by a doming in the coronal plane also, opposed to a slightly grooved (P2 /s/) or flat tongue shape. As a consequence, a narrow *palatal channel* is created along the tongue constriction, with the tightest point of constriction at its front end. The back cavity begins as far back as the post-palatal (P1) or velar (P2) region. This palatal channel sets /ç/ apart from the two other sibilants as it can be seen in the area functions.

The cross-sectional shape of tongue constriction is elliptic rather than circular in all of the fricatives, as seen in the coronal slices (Figures 7 and 8). Its cross-sectional area varies according to the speaker and the sibilant

type (ranging from about 5 to 25 mm<sup>2</sup>), and no systematic relationship could be established with the other articulatory or morphological characteristics.

In all the data, the jaw position is high, so that the vocal tract comprises a second constriction involving the incisors. The area of this second constriction is comparable to that of the tongue constriction, as seen on the area function (Figure 9). Because of this teeth constriction, the vocal tract portion that used to be designated as the *front cavity* appears as two distinct parts, the oral part (hereafter ‘front *oral cavity*’) comprising the sublingual portion (if any) as well as the cavity between the tongue constriction and the teeth constriction, and the lip cavity.

Although P1’s occlusion type has not been formally examined, he presents a very advanced position of the upper incisors with respect to the lower incisors (i.e. a large overjet). As a consequence, a small cavity is observed between the constriction formed by the upper incisors and the lower lip and a second constriction formed by the tongue and the alveolar ridge for /s/.

These teeth-inserted high resolution data show that the horizontal alignment of the upper and lower incisors vary according to the sibilants for both subjects, with a greater overjet (antero-posterior distance between the upper and lower incisors) for /ç/ than for /s/ and /ʃ/. It might be interpreted as the consequence of the articulatory requirements for tongue positioning in /ç/ with respect to the palate. We cannot exclude, however, the possibility that it is also related to noise source generation, where the lower teeth, as an obstacle, should be positioned differently with respect to the airstream that originate from constrictions of various shapes and distance depending on the fricative.

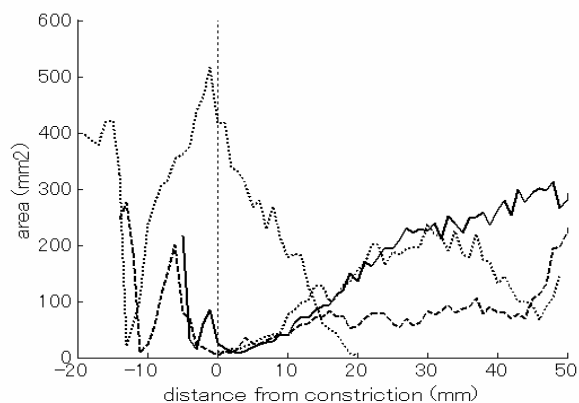


Figure 9. Area functions for P2; /s/ (plain line), /ç/ (broken line) and /ʃ/ (dotted line). The lip end is located at the left of the figure. The area functions are aligned at the tongue constriction (thin dotted line) For this speaker, only /ʃ/ comprised a sublingual cavity. This cavity is represented in the continuity of the front cavity, behind the constriction, in parallel with the back cavity. A second constriction, formed by the teeth, is observed at 3 mm (/s/), 11 mm (/ç/) and 13 mm (/ʃ/) in front of the tongue constriction, respectively.

#### 4. Formant-cavity affiliation in realistic models

A close observation of the vocal tract for Polish sibilant fricatives raises a question as to how the respective vocal tract cavities are reflected in peak positions of the frication noise. Also, if the formant-cavity affiliation were established, then it could lead us to account for the acoustic goal of the articulatory targets.

Although the acoustic structure of the variants of [s] merits to be examined in detail, this paper will focus on the non-anterior sibilants /ç/ and /ʃ/ because of the limitation of space. This contrast has been discussed in Halle and Stevens (1997), and retained our attention because their spectral difference does not show clearcut, even though their articulatory targets are distinctive as described in the previous section. Moreover, if native as well as non-native listeners are able to categorize the fricative from isolated frication noise (Lisker 2001, Nowak 2006), there should be consistent cues that are contrasting /ç/ from /ʃ/.

##### 4.1. Method

###### 4.1.1. Acoustic recordings

The acoustic signal was recorded in a soundproof room. The subjects P1 and P2 produced the isolated words *siadam* [çadam] and *szatan* [ʃatan], followed by the sustained fricatives [ç] and [ʃ] respectively, as a part of a larger corpus not analyzed here. This part of the corpus was produced in supine position in order to simulate the MRI condition.

The time-averaged spectra of the sustained fricatives are calculated by averaging the power spectra obtained with as many 5 ms hamming

windows (with 4,85 ms overlap) as necessary in order to cover the central 80 % portion of the fricatives. The averaging over many windows is based, within the source-filter theory framework, on the assumption that the noise source is random, thus potentially different in every window; however, the formant structure is assumed to be constant in a sustained utterance. So, larger the number of samples, least becomes the effect of accidental source fluctuation on the averaged spectrum. The 5 ms time window, which is shorter than usual, is however considered to be long enough for our purpose with regard to its frequency definition (200 Hz).

#### *4.1.2. Acoustic modeling*

For the acoustic modeling we used the VTF\_fric program (see section 3).

##### **Original models with artificially extended back cavity**

The vocal tract area functions were measured from the lips to about the velar region, as shown already in Figure 9, but the pharyngeal portion was discarded because of the poor contrast in the MRI data. Therefore, the original vocal-tract area function covers the front cavity, the constriction (including the palatal channel, if any), but only a short portion of the back cavity.

At this stage, the vocal tract area functions cannot be used in simulation, neither with a closed or open glottis because of the unrealistically short back cavity that results in erroneously higher resonance frequencies.

One possible solution to this problem is to eliminate the effect of sound reflections at the up-stream end of the short back cavity. An alternative approach consists in filtering out the back cavity resonances. This latter approach was chosen.

To do so, the rearmost section of the vocal tract was lengthened to about 1 m so that the resonances arising from this part of the model would be densely spaced. Actually, the spacing ranges from 161 to 182 Hz depending on the sibilant model. The densely distributed resonances of the back cavity therefore can easily be separated from the sparsely spaced resonances arising from the other cavities by cepstral smoothing.

Various source locations were used with the original models. Since pole frequencies do not change with respect to the source position or source type, and since the source position as well as its precise composition (monopole, dipole, etc.) are not known, we calculated all-pole functions by positioning the source at the glottis when the models did have no side branch.

Exciting the models at their rearmost point, which is not very realistic since the noise source is rather expected in the front, has the effect of emphasizing the relative amplitude of the resonances that arise from cavities located near the source. Therefore, the relative intensity of front cavity and palatal channel resonances might be inverted, while their respective frequencies are assumed to be correct.

Two other source locations, at the teeth and at the constriction, were used in the various models, including those which sublingual cavity was treated as a side branch. Our aim is not to examine the effect of all the exhaustive source locations, but to be able to observe the *frequency* of vocal tract resonances (invariant with respect to source location), so that they could be compared between original and truncated models (see below), and to interpret their cavity affiliation. By alternating source locations, we made certain not to miss a resonance, even if an antiresonance accidentally masked it in one of the source settings.

#### **Truncated models**

In fricatives, the tongue constriction is very tight (in order to meet the aeroacoustic requirements for the generation of a turbulent source), which minimizes the acoustic coupling between the anterior and posterior parts of the vocal tract separated by the constriction. Therefore, the acoustic properties of the anterior and posterior vocal tract portions should not differ much when estimated separately.

In order to examine the regions of the vocal tract that affect the spectral peaks, we run simulations with truncated vocal tract models, as in e.g., Fant (1970:182–184). *Anterior* and *posterior* vocal tract models were constructed by cutting the original area function at the tightest point of tongue constriction. The anterior models contain the front cavity, and the posterior models contain the palatal channel, if any. In posterior models, the rear portion was lengthened in the same manner as in the original models. In branching configurations, the front cavity models have the sublingual cavity as well as the front cavity connected one another, and their all-pole function was calculated by locating a source at the bottom of the sublingual cavity (the rearmost section of the model).

In addition, the anterior models were further truncated. In the previous section, where the vocal tract shape was described in detail, it was observed that the upper and lower incisors are put close together, so that the lip cavity is separated from the front oral cavity by a teeth constriction. Lip models were therefore created by truncating the vocal tract at the tightest point of teeth constriction.

## 4.2. Results

### 4.2.1. /c/

The actual noise spectra of the sustained utterances, as well as the cepstral-smoothed transfer functions of the various /c/ models, are shown in figure 10. Note that the natural noise spectrum is not equivalent to the calculated transfer functions. The former results from the combination of the source and transfer characteristics. As we mentioned in section 2, the source intensity is likely to be falling towards high frequencies, and this explains a difference of slope between the natural spectra and the simulation results. In addition, our original models do not comprise a realistic back cavity. However, in the natural spectrum, back cavity resonances are expected to appear as regularly spaced pole-zero bumps of small magnitude superimposed to the other resonances. This should partly explain the higher complexity of the natural spectra when compared to simulation results.

As shown at the left in Figure 10 for P1, the spectral prominence of the natural sound, which ranges from 2.5 to 7.5 kHz (with two characteristic peaks), coincides with the prominence with two peaks in the original model's transfer function. It is likely that these peaks come from the anterior and posterior vocal tracts, respectively, as the anterior (broken line) and posterior (dotted line) models' functions exhibit peaks near those frequencies. It can be assumed therefore that the natural peak originate in the clustering of the front cavity and palatal channel resonances.

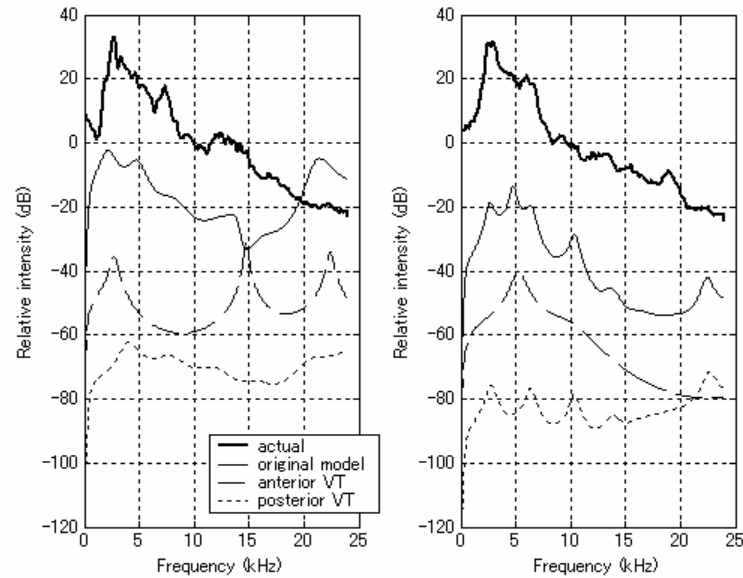


Figure 10. The actual averaged spectra (thick lines), and the simulated transfer functions (thin lines) for subject P1 (left) and P2 (right) /ç/. The original models (plain lines) as well as posterior vocal tract models (dashed lines) shown here were excited with a glottal source, whereas the anterior vocal tract models (broken lines) were excited with a source located at their backmost section.

Likewise, P2's actual sound's spectral prominence (2.5-7 kHz) with two major peaks corresponds well to the prominence made of three peaks in the original model's all-pole function. Contrary to P1, the second peak involves the anterior vocal tract, while the first and third peaks involve the posterior vocal tract. The reversed order of the front cavity and palatal channel resonances between the two speakers can be explained by a longer front cavity, which involves a sublingual cavity, and a shorter palatal channel in P1 with respect to P2, as can be seen in the MR images (Figures 7 and 8).

To summarize, the spectral prominence of /ç/ is likely to involve the combination of front cavity and palatal channel resonances (although the palatal channel resonance might be the lowest or the highest of the two, depending on the subject). Since the dimensions of the front oral cavity as well as that of the palatal channel can be adjusted independently, it can be

suggested that the subjects deliberately cluster those resonances in order to enhance the spectral prominence (which exhibits as much as 30 dB gain with respect to the low frequency dip), to achieve the spectral saliency.

#### 4.2.2. /s/

The results for /s/ are shown in Figure 11. The natural spectra of /s/ also exhibit a spectral prominence in which several peaks can be identified. The lowest peak (around 2 kHz) is somewhat lower in frequency than that of /ç/, while the second peak is the most prominent for P2. (In word initial position, contrary to the sustained utterances, the first peak was the most prominent also for P2.)

When the original model's transfer functions are compared to the natural spectra, the frequency of the spectral prominences roughly coincide with each other; however, there is a *missing* peak in the simulation results with respect to the natural spectra, i.e., the peak around 5 kHz that is the most prominent in P2's natural spectrum. Otherwise, the first and second peaks of the original model's function can be put in correspondence with the first and third peaks of the natural spectra, observed around 2 and 7.5 kHz, respectively.

In P1, the original model's peaks best coincide with the anterior vocal tract model's peaks. In P2, the posterior vocal tract contributes to some extent to the spectral structure of the original function. In general, in comparison to /ç/, the involvement of the posterior vocal tract seems to be rather small in /s/.



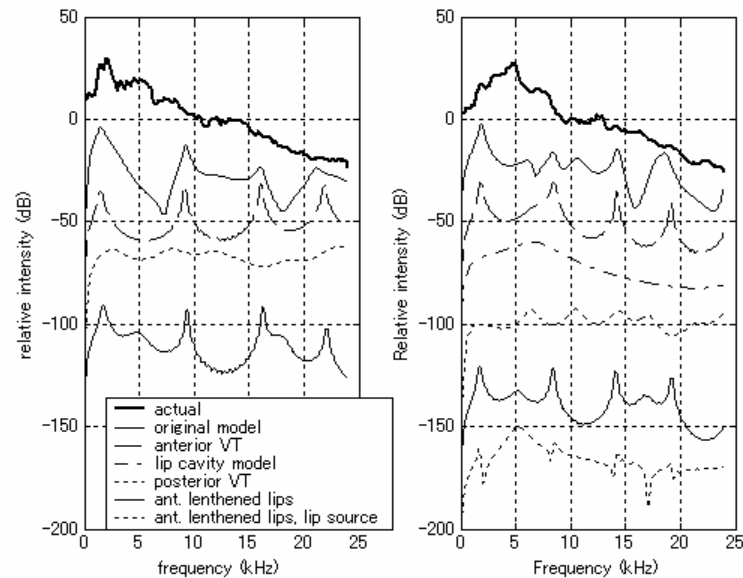


Figure 11. The actual averaged spectra of /s/ (thick lines), and the simulated transfer functions (thin lines) for subject P1 (left panel) and P2 (right panel). This figure shows the simulation results for the original models (upper plain lines) excited by a glottal source (P1) and a constriction source (P2), respectively. The results for anterior vocal tract models (broken lines) and lengthened lips anterior vocal tract models (bottom plain lines), as well as P2's lip cavity model (dotted-dashed line, right panel) are obtained with a source located at the backmost section. The simulation results for P2's lengthened anterior model excited with a source located within the lips are also shown (lower dotted line). The posterior vocal tract models (upper dotted lines) were excited at the glottis.

Looking more closely at P2's anterior vocal tract function, a broad resonance can be noticed, which forms a shoulder (around 7 kHz) at the left of the second peak. Because of its wide bandwidth, a lip resonance is suspected (*radiation load* is applied at the lip end). In order to verify this hypothesis, the acoustic properties of the lip model were examined. The dash-dotted line corresponds to the lip model's result. A single peak is observed near the frequency of the broad banded resonance observed in the anterior vocal tract transfer function, thus corroborating our hypothesis.

Therefore, we in fact have the right *number* of resonances corresponding to the three formants (around 2, 5 and 7 kHz, respectively) that constitute the natural utterance's spectral prominence.

We consider that the lip issue need to be examined more in detail, as a strong lip protrusion characterizes P2's MRI data of /s/, and we would like to be able to interpret its acoustic function. Actually, the lip gesture accompanies English and French /ʃ/, which are not far different from Polish /s/. It has been proposed (e.g. Stevens *et al.* 2004) that the acoustic finality of lip protrusion, at least in English /ʃ/, is to lower the front cavity resonance, thereby increasing the acoustic contrast with /s/. However, in our articulatory data, the lip cavity appears well separated structurally from the front oral cavity. Furthermore, our simulation results indicate that the lip cavity produces a formant by its own, distinct from those arising from the front oral cavity. These observations lead us to propose another interpretation. The function of lip protrusion would be, instead, to strengthen the spectral prominence by clustering the lowest lip resonance with a front oral cavity resonance, in order to enhance the spectral energy in this frequency region.

Is this plausible? We have mentioned that the area of some lengths of the lips could not be measured because of their lateral curvature. In P2's area function, the measured lip cavity is 6 mm long. On his mid-sagittal profile, however, it clearly exceeds 1 cm. The frequency of the lip cavity resonances are therefore largely overestimated in our simulation. If the anterior vocal tract model is altered, with the lip cavity lengthened by 5 mm, the lip resonances move down, as shown in the anterior vocal tract all-pole function (lower plain line). When the source is put within the lip cavity, the relative intensity of the first lip resonance get much closer to the actual spectrum, as shown in the lowest dotted curve. Note that it is not unrealistic to assume a lip noise source, created by the air jet exiting from the teeth constriction and directed against the lower lip.

A similar correction can be applied to P1's /s/. For this speaker, no lip cavity could be measured, and thus the models contain no lip cavity. This is clearly far from the reality even though no lip protrusion is observed in /s/ for this speaker. If an artificial lip cavity of 1 cm of length and 4 cm<sup>2</sup> of area is appended to P1's front cavity model, an additional peak appears around 5 kHz, which nicely match the second formant of the actual spectrum, without significantly affecting the other peaks.

To summarize, /s/'s spectral prominence involves mainly the front cavity resonances (corresponding to the actual F1 and F3). In addition, it is

very likely that the lip cavity is responsible for the second fricative formant of /ʃ/.

#### 4.3. Discussion

From these results, the following observations can be made:

- (1) The first fricative formant for /ç/ tends to be higher in frequency than that of /ʃ/. The size of the front oral cavity (larger for the latter) is largely responsible for this difference.
- (2) In both /ç/ and /ʃ/, the spectral prominence is made of a set of resonances arising from distinct cavities. In our results, the first front oral cavity resonance is clustered with the first palatal channel resonance in /ç/, whereas the lip resonance is clustered with the front oral cavity resonances in /ʃ/.

If the speakers aim at enhancing the spectral saliency of these sounds, they will tend to bring the resonances arising from different cavities close one to another, by adjusting the lengths of these cavities. If the clustering patterns described in (2) could be generalized, then we could predict that:

- (3) In /ç/, the front oral cavity resonance can be brought very close to the palatal channel resonance (this is still low enough so that this sound is distinct from /s/). Therefore, /ç/'s prominence will tend to be compact.
- (4) In /ʃ/, if the goal is to make a very low fricative formant, the front oral cavity is the only vocal tract region that leads to this goal. However, due to the physical limitations of lip protrusion, the lip resonance cannot be grouped with the first front cavity resonance as close as the palatal channel resonance does in /ç/. Moreover, the lip resonance possesses a wider bandwidth than the palatal channel resonance. Therefore, /ʃ/'s spectral prominence will tend to be more diffuse.

These predictions seem to be verified by our acoustic data. It is however difficult to make a generalization until the acoustic properties of frication noise are quantitatively examined by means of ensemble-averaged or long-term spectra from a larger set of Polish speakers.

We would like to add a few words about the coupling of the back cavity, which was assumed to be partly responsible for the spectral difference between the Polish sibilants /ç/ and /ʃ/ in Halle and Stevens (1997:190). Halle and Stevens (1997:189) argued that the long and thin palatal channel prevents the back cavity resonances from coupling effectively in /ç/, in contrast to /ʃ/. A back cavity resonance would have been responsible for

*/s/*'s characteristic peak in the vowel's F2 region. However, it cannot easily be conceived that a pole-zero pair of the back cavity would be able to shape significantly the noise spectrum of */s/*, given that the constriction is very tight, measuring about 0.1 to 0.15 cm<sup>2</sup> in our MRI data (see Badin 1989: 52-53). In these conditions, not only closely bound back cavity poles and zeros are expected to be considerably attenuating each other, but in addition, the back cavity would be very weakly excited by the noise source. Indeed, the long-term spectral structures of sustained utterances of */s/* do not exhibit such kind of regular formant-antiformant pairing big enough to shapen the overall spectral shape.

Instead, in the light of our results, it can be suggested that the larger number of peaks observed in the lower frequencies for */s/* are mainly due to the longer front oral cavity (including the sublingual cavity), which leads to lower resonances, combined with the lip resonance.

## 6 Conclusion

This study aimed at interpreting analytically the spectral structure of sibilant fricatives. First, the acoustic properties of a simple sibilant-like configuration were examined through acoustic simulations by varying the straight/branching configuration and the source location. The spectral events characteristic of sibilant fricatives were then summarized.

Second, the articulation of Polish sibilants */s*, *ç* and *ʂ* were examined in detail using the high-resolution MRI data from two subjects. The data showed the existence of a second tight constriction located at the incisors, in addition to the constriction involving the tongue. The effect of this second constriction on a possible lip noise source remains to be investigated.

Finally, the formant-cavity affiliation for these subjects' non-anterior sibilants */ç* and *ʂ* was discussed, by simulating the acoustic properties of original and truncated models constructed from the MRI data. The main outcome of this experiment is to have pointed out that in both */ç* and *ʂ*, the spectral prominence is likely to be made up of a bunch of resonances arising from different vocal tract cavities: front oral cavity and palatal channel for */ç*; and front oral cavity and lip cavity for */ʂ*. The spectral consequence of the presence/absence of a palatalized articulation is a lower (*/ʂ*) or higher (*/ç*) frequency of the spectral prominence, as well as its relative diffuseness (*/ʂ*) or compactness (*/ç*). Contrary to Halle and Stevens's (1997) proposal, our articulatory data as well as simulation

results do not support a significant coupling of back cavity resonances in /ʃ/. Instead, the spectral difference between /ç/ and /ʃ/ is better explained by the difference in size of the front oral cavity, combined with other resonances (palatal channel and lip cavity resonances, respectively).

### **Acknowledgement**

This work was supported, in part, by the ASPI project, EC 6<sup>th</sup> Framework Program n. 021324. We are grateful to the Polish subjects, who provided sustained efforts in order to permit the acquisition of high quality data, and also to Yasuhiro Shimada, Ichiro Fujimoto, and Yuko Shakudo of the ATR Brain Activity Imaging Center for their technical expertise.

### **References**

Badin, Pierre

1989 Acoustics of voiceless fricatives: production theory and data. *STL-QPSR* 3/1989: 33–55.

Curle

1955 The influence of solid boundary upon aerodynamic sound. *Proc. Roy. Soc. Lond. A* 231: 505–514.

Fant, Gunnar

1970 Reprint. *Acoustic theory of speech production*. Second printing. The Hague and Paris: Mouton. Original edition: The Hague: Mouton, 1960.

Flanagan, James L.

1972 Second edition. *Speech analysis, synthesis and perception*. Berlin, Heidelberg and New York: Springer-Verlag. First edition: Berlin and Heidelberg, Springer-Verlag, 1965.

Halle, Morris and Kenneth N. Stevens

1997 Reprint. The postalveolar fricatives of Polish. In S. Kiritani, H. Hirose and H. Fujisaki (eds.) *Speech production and language: in honor of Osamu Fujimura*. Berlin: Mouton de Gruyter: 177–193. Originally in *Speech communication group working papers* 7, Research laboratory of electronics, MIT, 1991: 77–94.

Harris, Katherine S.

1958 Cues for the discrimination of American English fricatives in spoken syllables. *Language and speech* vol. 1: 1–17.

Jassem, Wiktor

1995 The acoustic parameters of Polish voiceless fricatives: an analysis of variance. *Phonetica* vol. 52: 251–258.

Lee, Sook-hyang, Mary E. Beckman and Michel Jackson

1994 Jaw target for strident fricatives. *Proc. ICSLP 2004*: 37–40. Yokohama, Japan.

Lisker, Leigh

2001 Hearing the Polish sibilants [s ś š]: phonetic and auditory judgements. *Travaux du cercle de linguistique de Copenhague* vol. XXXI: 226–238.

Maddieson, Ian and Kristin Precoda

1989 Updating UPSID. *UCLA working papers in phonetics* 74: 104–111.

Maeda, Shinji

1982 A digital simulation method of the vocal-tract system, *Speech Communication* 1: 199–229.

Mooshammer, Christina, Philip Hoole and Anja Geumann

2007 Jaw and order. *Language and Speech* vol. 50, no. 2: 145–176.

Narayanan, Shrikanth and Abeer Alwan

2000 Noise source models for fricative consonants. *IEEE transactions on speech and audio processing* vol. 8, no. 2: 328–344.

Narayanan, Shrikanth S., Abeer A. Alwan, and Katherine Haker

1995 An articulatory study of fricative consonants using magnetic resonance imaging. *Journal of the acoustical society of America* 98 (3): 1325–1347.

Nowak, Paweł M.

2006 The role of vowel transitions and frication noise in the perception of Polish sibilants. *Journal of phonetics* 34: 139–152.

Pastel, Leah M. P.

1987 Turbulent noise sources in vocal tract models. M.S. dissertation. Massachusetts Institute of Technology.

Shadle, Christine H.

1985 The acoustics of fricative consonants. PhD Dissertation, MIT.

1991 The effect of geometry on source mechanisms of fricative consonants. *Journal of phonetics* 19: 409–424.

Stevens, Kenneth N.

1998 *Acoustic phonetics*. Cambridge, Massachusetts and London, England: MIT press.

Stevens, Kenneth N., Zhiqiang Li, Chao-Yang Lee and Samuel Keyser

2004 A note on Mandarin fricatives and enhancement. In G. Fant, H. Fujisaki, J. Cao and Y. Xu (eds.), *Foreign language teaching and research press*, Beijing, 393–403.

Takano, Sayoko, Kiyoshi Honda, Shinobu Masaki, Yasuhiro Shimada and Ichiro Fujimoto

2003 Translation and rotation of the cricothyroid joint revealed by phonation-synchronized high-resolution MRI. *Proc. EuroSpeech 2003*: 2397–2400, Geneva.

Takemoto, Hironori, Tatsuya Kitamura, Hironori Nishimoto and Kiyoshi Honda

2004 A method of tooth superimposition on MRI data for accurate measurement of vocal tract shape and dimensions. *Acoustical Science and Technology*, 25 (6): 468–474.

Toda, Martine

Forthcoming. Etude articulatoire et acoustique des fricatives sibilantes. Doctoral dissertation, Université Paris III, France.

Wierzchowska, B

1965 *Wymowa Polska*. Warszawa: Panstwowe Zaklady Wydawnictwo Szkolnych.

Żygis, Marzena, and Silke Hamann

2003 .Perceptual and acoustic cues of Polish coronal fricatives. *Proc.15<sup>th</sup> ICPHS*: 395–398, Barcelona.