Acoustics of voiceless fricatives: Production theory and data

Badin, P.

journal: STL-QPSR
volume: 30
number: 3
year: 1989
pages: 033-055

http://www.speech.kth.se/qpsr
ACOUSTICS OF VOICELESS FRICATIVES: PRODUCTION THEORY AND DATA

Pierre Badin*

Abstract

The aerodynamic and acoustic phenomena involved in the production of fricative consonants are far from being completely understood. The overall level and spectral characteristics of the radiated sound pressure are known to primarily depend on the aerodynamic state of the vocal tract, i.e., the pseudostatic pressure drop across the main oral constriction and the cross-sectional area of the constriction. The first part of this paper describes an attempt to establish some quantitative relationships, based upon experimental measurements carried out on a human subject, both for sustained and dynamic voiceless fricatives. The second part of the paper focuses on the acoustic modelling of the vocal tract in the frequency domain; it describes our attempts to match some measured spectra with synthetic spectra computed from simplified area functions, and the influence of different parameters on these spectra. We show, among other things, that the obstacle effect can be explained in terms of filter characteristics.

INTRODUCTION

The acoustic and aerodynamic phenomena involved in the production of fricative consonants are far from being completely understood. In order to improve our modelling, we have tried to confront the current knowledge in that field with new experimental data. The production of fricative consonants involves a constriction in the vocal tract, where the local particle velocity is increased, due to the small cross-sectional area, and where the flow shifts from laminar to turbulent; this creates the equivalent of a serial noise pressure source. We intend to relate the overall level and spectral characteristics of this pressure source and of the radiated sound pressure to the aerodynamic state of the vocal tract. On the other hand, we want to investigate more thoroughly the properties of the vocal tract acoustic transfer function in the frequency domain in relation to source location and impedance functions of simplified cavity structures. Thus, in the first part of the paper, we deal with characteristics related to acoustics and aerodynamics, and in the second part we describe simulation studies more directly related to overall source-filter production models.

1. ACOUSTICS

1.1 Previous studies

Before describing the experimental measurements that we have performed, we review previous theoretical and experimental studies in this field. We are interested in estimating both overall levels and spectral characteristics of the noise source. There are, so far, no direct experimental means for measuring the source pressure itself; this has to be inferred through the sound pressure radiated at the lips.

* Institut de la Communication Parlée (UA CNRS), INPG - Univ. Stendhal,
46, av. F. Viallet, F-38031 Grenoble Cedex, FRANCE
Basically, the aerodynamic modelling of fricatives aims at defining the characteristics of the noise source as a function of the vocal tract aerodynamic variables, i.e., the pseudo-static pressure drop $\Delta p$ and flow $U$ across the main oral constriction, and the constriction cross-sectional area $A_c$.

### 1.1.1 Source functions and overall level

Various relations have been proposed for overall fricative source pressure $P_s$ or radiated sound pressure $P_r$ as a function of the pressure drop across the constriction $\Delta p$ and of the constriction area $A_c$.

Fant (1960) adopted a serial noise pressure source proportional to the particle velocity squared, $V_c^2$, and thus to the pressure drop $\Delta p$ in the constriction approximated by the Bernouilli equation $\Delta p = kpV_c^2/2$. This was applied to the reconstruction of spectra from area functions of Russian consonants. The source was assumed to be independent of the constriction area $A_c$.

Earlier experimental measurements (Meyer-Eppler, 1953) established that the radiated sound pressure was proportional to the Reynold's number $Re = V_c d / \nu$ above a certain threshold and thus to $V_c^2 A_c$.

Stevens (1971) inferred from theoretical and experimental results on spoiler-generated noises in ducts that the source pressure level is proportional to $\Delta p^{3/2} A_c^{1/2}$, leading to the relation: $P_s \sim V_c^3 \sqrt{A_c}$ (Stevens, 1987).

It is of interest to confront these relations with equations that could be derived from the more recent study on noise generated by spoilers in ducts by Nelson & Morfey (1981). For frequencies below the higher mode cutoff frequency of the vocal tract, which presumably lies in the range 4-8 kHz, the radiated power in a given frequency band may be expressed as (Nelson & Morfey, p. 276):

$$W_{\Delta f} = \frac{1}{4} \frac{A}{\rho c} K^2(S_t) \Delta p^2,$$  \hspace{1cm} (1)

where $K(S_t)$ has been experimentally shown to be inversely proportional to the normalized frequency $S_t = fd/V_c$, called the Strouhal number, $d$ is a characteristic dimension of the constriction (Nelson & Morfey, p. 283) leading to $d = \sqrt{A_c \pi}$, and $A$ is the duct cross-sectional area.

Assuming a plane wave source loaded by an infinite duct, and thus by an impedance $Z_L = \rho c A$, we can write:

$$W_{\Delta f} = P_s^2 / Z_L,$$  \hspace{1cm} (2)

and thus:

$$P_s = \left( \frac{1}{2} K_0 \cdot \frac{V_c}{d} \cdot \Delta p \right) \frac{1}{f},$$  \hspace{1cm} (3)

which finally leads to $P_s \sim V_c^3 \sqrt{A_c}$. 

34
In 1967, Hixon, Minifie, & Tait measured Intra Oral air Pressure (IOP), volume rate of airflow, and noise sound pressure level (SPL) during [ʃ] and [s] productions of two speakers. They found a sound pressure (SP) approximately proportional to the 1.2-1.4 powers of IOP and a flow proportional to the 0.5-0.6 powers of IOP for both [ʃ] and [s]. This corresponds to overall sound pressure proportionalities of the order of \( V_c^{2.4} \) to \( V_c^{2.8} \).

We should finally mention Shadle’s important contribution to the field (1985). She made experiments similar to those of Hixon & al. on highly simplified mechanical models of fricative configurations, and tried to “characterize the change in [radiated] sound pressure due to flow velocity [...] by the power exponent \( n \), defined by \( P \sim V^n \), where \( P \) is the total sound power generated by a flow of velocity \( V \).” She refers to the quantity \( n/2 \) as the pressure exponent. She shows that for the obstacle case, i.e. the case where the source is mainly located at an obstacle (usually the teeth) hit by the air jet produced by the constriction, the exponent \( n/2 \) ranges between 2.5 and 2.7, whereas for the no-obstacle case, it ranges between 2.7 and 3.7.

For more clarity, we will in this paper refer to the parameter \( p \) in the relation \( P_r \sim (\Delta p)^p \) as the "IOP" exponent, to the parameter \( q \) in \( P_r \sim (A_c)^q \) as the "Area" exponent, and to the parameter \( r \) in \( P_r \sim (V_c)^r \) as the "velocity" exponent. Observe that this terminology is slightly different from that of Shadle.

In 1988, Shadle performed direct measurements on mechanical models reproducing fairly accurately some fricative configurations. With a small probe microphone, she measured noise pressure spectra in the regions of possible source locations along the models. However, these are not immediately interpretable as serial source data.

1.1.2 Spectrum
A more detailed modelling of the noise source has to consider the spectrum shape as well, and its non-uniform growth with \( V_c \).

Based again on results on spoilers, Stevens (1971) assumes that the source spectrum has a broad maximum around the Strouhal numbers of 0.2.

Shadle (1985) has determined the spectrum of the source from the radiated sound in mechanical model experiments by inverse filtering with the known transfer function for an assumed location of the source. For obstacle cases, the source spectra slopes increase with frequency, but do not exhibit any kind of broad peak. The source spectrum is not explicitly determined for the no-obstacle case.

She also determined "velocity" exponents for different frequency bands to estimate the non-uniform growth of the spectral levels with \( V_c \) (see further Fig. 6).

1.2 Our experimental measurements
We describe in this section the measurements carried out on human sustained and dynamic fricatives that we have performed in order to infer the best modelling for the relations between the aerodynamic parameters and the overall radiated sound level and spectral characteristics.

1.2.1 General experimental setting and procedures
We have replicated and extended the experiments of Hixon & al. (1967). An ideal set-up would have allowed a simultaneous measurement of pseudo-static pressure drop at the constriction, flow and overall radiated sound pressure level. However, we assume that the pressure drop at the constriction is almost identical to the Intra Oral Pressure (IOP) since the front cavity has a negligible aerodynamic resistance. The IOP and the flow can be measured with any kind of face mask-pneumotachograph, whilst radiated sound pressure will be disturbed (Hixon, 1966, p. 171; Badin, Karlsson, & Hertegård, forthcoming). We had thus to split the
Fig. 1. Example of measured fricative spectra for [l], [s], and [f].
Fig. 2. Distribution of the minimum constriction area for [ʃ], [s] and [f] versus IOP (Mask series).
In both series, the French subject PB was seated in an anechoic room; a Brüel & Kjær microphone was placed at a distance of 20 cm in front of his mouth, slightly out of axe to avoid wind effects. The signal was recorded on a Digital Audio Tape recorder. A calibration procedure allowed us to keep the absolute SPL values.

During the first series, the subject wore a Rothenberg mask, which is a circumferentially vented pneumotachograph, especially designed to minimize any acoustic distortion of the speech signal (Rothenberg, 1973). In the commercially available version of the mask, beside the flow transducer, there is provision for a pressure transducer; it was connected to the cavity behind the main oral constriction through a small polyethylene tube inserted through the lips at the mouth corner and running on one side of the mouth between the cheek and the gum, thus allowing the IOP measure. The IOP and flow signals were recorded on two channels of a Frequency Modulation recorder. The signals were then digitized at 20 kHz sampling rate for the radiated sound pressure, and at 10 kHz for IOP and flow, together with synchronization markers.

During the series of experiments without mask, only the pressure transducer was used, and thus only the IOP signal recorded.

For both sustained and dynamic fricatives, the corpus was limited to three voiceless fricative consonants [f], [s] and [ʃ], because it was impossible to record the IOP for more retracted articulations with this method.

### 1.2.2 Sustained fricatives

Our first approach was to restrict ourselves to sustained fricatives in order to get simpler experiments.

**Data recording and processing**

The subject was instructed to sustain the syllables [aʃ], [as], and [af], keeping the fricative segment for at least 500 ms, at different levels from the lowest effort to the highest possible (however, our pressure transducer was limited to 10 cm H2O).

The digital signals were then computer-edited to extract 500 ms segments in the most stable part of each sound: the absolute SPL, the average IOP and the average flow were estimated over the total segments.

The spectra of the sound pressure signal recorded for each of the items in the experiment without mask were estimated from the Long-Time-Average-Spectrum method. These LTAS were computed by averaging 80 FFT spectra from overlapping Hamming windowed segments over 500 ms of signal. The analysis bandwidth was 80 Hz, corresponding to a 12.5 ms window, and to 512 points FFT. Fig. 1 shows an example for each of the three fricative classes.

**Aerodynamic parameters**

From the IOP Δp and the flow \( U = V_c A_c \), we can estimate the minimum constriction area from the well known "orifice equation" (see, for instance, Warren & DuBois, 1964 and Hixon, 1966) derived from the Bernouilli equation \( \Delta p = k \rho V_c^2 / 2 \) and \( k = 1 \):

\[
A_c = \frac{U}{\sqrt{2 \cdot \Delta p / \rho}}
\]  

(4)

Observe that these areas are "equivalent minimum aerodynamic areas and not the actual physiologic areas associated with the maximum constriction" (Hixon, 1966).

Fig. 2 shows the distribution of the minimum constriction area for the three fricatives. The area for [ʃ] varies between 0.1 and 0.4 cm². The subject seems to use two different strategies to produce a low IOP: either a small constriction area, and thus a low flow, or a bigger area and a higher flow: this might be related to different articulatory positions. The minimum constriction area for [s] is approximately independent of the IOP, around 0.1 cm². For [ʃ], the av-
average is also around 0.1 cm², with a slightly higher scattering.

Fig. 3 displays the flow versus IOP data. The slopes of the linear regression lines for [s] and [f] are both 0.6; for [J] the scattering of constriction areas leads to a slope of 0.2, that can be brought up to 0.4 if we remove the points with an area greater than 0.2 cm² (see further column 4 in Table I). If the minimal constriction area were constant for different IOP, this slope would be 0.5. The departure from this law shows a tendency for the area to increase slightly with the IOP. This was also noted by Hixon & al. (1967) for their subjects.

In general, all our results check fairly well with those of Hixon & al., if we observe that they used a slightly different version of the Bernouilli equation, providing 40% lower areas.

**Overall level of radiated sound pressure**

Fig. 4 shows overall SPL versus IOP for the series without mask. Since the IOP is expressed on a logarithmic scale, the slope of the linear regression lines are direct expressions of the "IOP" exponents defined above. These are given in Table I.

In order to determine from the data for the mask series the "IOP" and "Area" exponents, i.e., the parameters $p$ and $q$ in the relation $Pr = (\Delta p)^p A_C^q$, we have used multiple linear regression, choosing SPL as the dependent variable, and $\Delta p$ and $A_C$, after conversion to a logarithmic scale, as independent variables (see Table I).

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>[J]</td>
<td>1.00</td>
<td>1.32</td>
<td>1.00</td>
<td>0.20</td>
<td>1.27</td>
<td>1.4</td>
<td>1.2</td>
</tr>
<tr>
<td>[J]*</td>
<td>1.32</td>
<td>1.34</td>
<td>0.14</td>
<td>0.43</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>[s]</td>
<td>1.26</td>
<td>1.18</td>
<td>0.72</td>
<td>0.60</td>
<td>1.42</td>
<td>1.3</td>
<td>1.3</td>
</tr>
<tr>
<td>[f]</td>
<td>1.04</td>
<td>0.97</td>
<td>0.41</td>
<td>0.60</td>
<td>1.31</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

(1) "IOP" exponents for PB (Mask series, assumes no $A_C$ dependency)
(2) "IOP" and (3) "Area" exponents for PB (Mask series)
(4) "Flow versus IOP" exponents for PB (Mask series)
(5) "IOP" exponents for PB (No-Mask series)
(6) "IOP" exponents for FDM (Hixon & al. data)
(7) "IOP" exponents for TJH (Hixon & al. data)

* The second [J] is similar to the first one, but all the data for $A_C>0.2cm^2$ are removed.

**Table I. "IOP" and "Area" exponents for human sustained fricatives.**

**Spectrum of radiated sound pressure**

Fig. 5 shows the evolution of the spectrum for [J] with the overall SPL. We clearly see that the high frequency part of the spectrum increases with overall SPL more rapidly than the low frequency part. In order to specify quantitatively this change, we have computed the SPL in 10 frequency bands of approximatively 1 kHz width, and for each band estimated the "velocity" exponent, assuming an approximatively constant constriction area, i.e., a flow proportional to $\sqrt{\Delta p}$. The results are shown in Fig. 6. The fricative [J] exhibits the greatest spectrum tilt change, whereas [f] shows the smallest, and less regular changes. This appears to reflect variations in the source spectrum rather than in the transfer function.
Flow versus IOP for [j] (Correlation = 0.55)

Intra Oral Pressure (dB Re 1. cmH₂O)

Flow versus IOP for [j]

Intra Oral Pressure (dB Re 1. cmH₂O)

(Ac < 0.2 cm²)

Flow versus IOP for [k]

Intra Oral Pressure (dB Re 1. cmH₂O)

Correlation = 0.90

Flow versus IOP for [s]

Intra Oral Pressure (dB Re 1. cmH₂O)

Correlation = 0.96

Flow versus IOP for [f]

Intra Oral Pressure (dB Re 1. cmH₂O)

Correlation = 0.82

Fig. 3. Flow versus IOP for [j], [s] and [f] (Mask series).
Overall SPL versus IOP for [j]

Correlation = 0.94

Overall SPL versus IOP for [s]

Correlation = 0.81

Overall SPL versus IOP for [f]

Correlation = 0.91

Fig. 4. Overall SPL versus IOP for [j], [s] and [f] (No mask series).
Fig. 5. Example of evolution of [ᵢ] spectra for increasing overall SPL (No mask series). (Overall SPL increases from 53 dB to 76 dB by 3 dB steps from bottom to top - Curve spacing = 10 dB.)

Fig. 6. "Flow" exponents for human fricatives and mechanical models. [ᵢ] = "+", [s] = "o", [f] = "#", Shadle's obstacle case with front cavity length 3.2cm = "x". (The "zero" frequency corresponds to overall SPL exponents.)
1.2.3 Fricative dynamics

Data recording and processing
In order to reconfirm our results on sustained fricatives for more natural articulatory conditions, and also to gain insight into the dynamics of fricative, we have recorded another small fricative corpus on the same subject: the previously described experimental setting was used for a number of [Ca] tokens, where C was either [ʃ], [s], or [ʃ].

For each logarom, the three signals (sound pressure, IOP, and flow) were processed in the following way. The voiceless parts of [Ca] and [aC] were edited in order to measure the SPL level of the noise component only, excluding any possible voiced component in the transition between the fricative and the vowel, we excluded also signals of which the level was more than 10 dB below the background noise level (36 dB). Then the dynamic evolution of the SPL was estimated from successive 5 ms windows, and smoothed out with an 80 Hz low-pass filter. The IOP and flow signals were also low-pass filtered at 80 Hz, and the constriction area evolution was computed from those according to the orifice equation. Finally, the three signals were undersampled at 1 kHz, to produce, for each fricative transition, sets of measured SPL, Δp, and A_c data of 100-300 points.

Aerodynamic parameters
Fig. 7 shows, on logarithmic scales, the example of the dynamic evolutions of SPL, IOP, flow, and area for [ʃæl]. The segments chosen for the analysis are shown by vertical bars. This example, typical of our corpus, shows the maximum constriction in the middle of the two fricatives. It should be noted that the estimation of the constriction area holds for consonants only, and that the values obtained for the vowel parts should therefore be ignored.

Fig. 7. Measured SPL, IOP, flow and derived constriction area for [ʃæl].
Overall level of radiated sound pressure

The sets of (SPL, Δp, A_c) data were used to perform a multiple regression similar to the one mentioned above for sustained fricatives, in order to determine the "IOP" and "Area" exponents individually for each transition. The exponents were also estimated globally for each fricative class. Table II gives these exponents, as well as those derived from the SPL/IOP data neglecting the influence of A_c, for comparison purposes.

Table II. "IOP" and "Area" exponents for human dynamic fricatives.

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>[f]</td>
<td>1.34</td>
<td>(0.98/1.66)</td>
<td>0.42</td>
</tr>
<tr>
<td>[s]</td>
<td>1.43</td>
<td>(0.81/2.09)</td>
<td>0.29</td>
</tr>
<tr>
<td>[f]</td>
<td>0.77</td>
<td>(0.62/1.68)</td>
<td>0.07</td>
</tr>
</tbody>
</table>

(1) "IOP"
(2) "Area" exponents for PB;
(3) "IOP" exponent for PB (no A_c dependency assumed).

The exponents are estimated over the whole mask data set; in parenthesis are given the minimum and maximum values of the individual estimations.

Spectrum of radiated sound pressure

In order to check if we could find in transitions any traces of the spectrum tilt changes with overall SPL, we generated spectrograms of the radiated sound pressure, and for the transitional segments (beginning and end of the fricative) computed spectral sections every 2 ms. It appears clearly (see Fig. 8) that the spectral tilt changes during the transitions, and that high frequencies appear after the low frequencies onset during the transition into the fricative and disappear before the low frequency offset during the transition out of the fricative.

Fig. 8. Spectrogram and spectral sections for [a]. Left hand side: spectrogram; right hand side: spectral sections every 2 ms (from 4 to 34 ms). (In both cases, the analysis bandwidth is 250 Hz). (The dashed lines highlight the limits of the noise component).
1.2.4 Discussion

Before comparing the data available in the literature with our own data, we should make three remarks.

The assumption of a direct proportionality between the overall source level and the overall radiated sound pressure level holds only for source signals having a uniform spectral dependency on $\Delta p$ and $A_c$, and approximatively for fricatives having one main resonance only. This distinction between the pressure source and the radiated sound pressure has generally not been made as judged from the literature; moreover, we have shown that this proportionality is in fact frequency dependent. However, with only one exception (Shadle, 1988), all the experimental data in the literature, as well as ours, have been obtained from measurements of the radiated sound pressure, and not of the source itself. We will thus discuss the comparisons on the basis of radiated sound pressure.

From our data we have derived "IOP" and "Area" exponents; Shadle (1985) has derived "velocity" exponents. It is clear that the measured IOP includes the Bernouilli pressure drop as well as the pressure drops due to frictional and laminar losses; however, since our simulations have shown that, in our experimental conditions, the laminar losses are insignificant and the frictional losses at most 10% of the total IOP, we are entitled to assume in the following that $\Delta p \sim V_c^2$, neglecting the loss terms. The "velocity" exponents are therefore twice the "IOP" exponents.

The comparison between columns (2) and (5) in Table I shows that the "IOP" exponents obtained from measurement with or without the mask do not differ too much, for [I] and [s], at least. We thus use the exponents derived from the data with mask in order to be able to assess the $A_c$ dependency as well.

The comparison between Tables I and II shows that there is no major incoherence between the two sets of measurements. We can not expect much closer results, since the data present a fair amount of scattering; for instance, the correlation factors corresponding to the regression of the SPL against IOP and $A_c$ are 0.85, 0.77, and 0.61 (cf. also maxima and minima in Table II) respectively for [I], [s], and [f].

In order to have a more concrete evaluation of the fit of these coefficients with measured values, we have resynthesized, for each transition, the SPL dynamic evolution as a function of $\Delta p$ and $A_c$, from the equation:

$$\text{SPL} = G_0 + p \cdot 20 \log_{10}(\Delta p) + q \cdot 20 \log_{10}(A_c),$$

(5)

using for $G_0$, $p$, and $q$ three different sets of values obtained by multiple regression in different conditions: (1) of SPL against $\Delta p$ and $A_c$ for each individual transition, (2) of SPL against $\Delta p$ and $A_c$, globally for each fricative class, and (3) of SPL against $\Delta p$ only, for each fricative class. Fig. 9 shows an example of comparison between the measured SPL and different synthesized versions. A visual inspection of such curves for all transitions shows that, apart from the residual oscillations of the measured curve (related to the short-term fluctuation of the noise amplitude), the maximum error between the measured and synthesized version is of the order of magnitude of 5-6 dB. Moreover, we have found that in most cases for [I] and [s], the fit with an $A_c$ dependency is better than the one without; this reconfirms this dependency.

The "velocity" exponents derived from our data, as well as the corresponding data from Shadle for obstacle and non-obstacle configurations (p. 121 and p. 127), are displayed in Fig. 1 (see the zero frequency band). The agreement with Shadle’s data is fairly good, and we see that our [I] and [s] have "velocity" exponents close to the value 2.5 obtained by Shadle for her obstacle configuration with 3.2 cm long front cavity.
From the comparison, it also comes out that our [f] behaves rather as an obstacle configuration than a non-obstacle one. Our subject reported that during the production of [f], his upper incisors were in contact not with the lower lip edge, but with the internal surface of it: it is likely that the air jet produced by that constriction hits the lower lip surface, or at least follows it. This could explain the exponent for [f].

We finally recall that our results are extremely close those of Hixon (1966).

1.2.5 Conclusions

The previous analysis shows that, in all cases, the "velocity" exponents that we have derived lie between 2 and 3, for [f] and [s], with "Area" exponents between 0 and 0.5. These values correspond to a situation intermediate between the equation proposed by Stevens and the equation proposed by Fant; the other proposals mentioned in Section 1.1.1 can be discarded. We suggest thus to use for [f] and [s] p=1.3 and q=0.3. The consonant [f] behaves in a fairly different way, and more suitable values seem to be p=0.8 and q=0.2.

These were preliminary experiments: only one subject on a very limited set of configurations has been examined. The experimental method has provided interesting data: we now need to investigate a larger corpus with more subjects to infer more general rules, especially for the spectral changes. We should also substitute the IOP probe by one inserted through the nasal cavity, in order to be able to treat also the more retracted consonants. It will also be of great interest to check the validity of these equations in a synthesis scheme, and to derive the proper spectrum dependant corrections to apply these equations to the source itself.

2. SOURCE-FILTER MODELLING

This section deals with the source-filter modelling of the fricative production in the acoustic domain, that is, the simulation of the acoustic transfer function of the vocal tract for fricatives.
2.1 Acoustic and pseudo-static aerodynamic modelling

The first stage of our frequency domain acoustic modelling of the vocal tract is a pseudo-static aerodynamic model that defines some boundary conditions (glottis and constriction resistances) from aerodynamic parameters (subglottal pressure, glottis opening and constriction area).

From an aerodynamic point of view, the vocal tract can be considered as a tube with two lumped constrictions, namely the glottis and the main oral constriction, driven by a pseudo-static subglottal pressure source $P_{sgl}$.

The glottis is modelled in a non-vibratory state, according to Fant (1960, p. 267), as slit of depth $l$ (0.3 cm) in the flow direction, and of width $w$. One can consider that the static pressure drop across the glottis is defined by:

$$P_{g\ stat} = P_{g\ lam} + P_{g\ ber},$$

where

$$P_{g\ lam} = (12 \ \mu l/w^2) \cdot (U/A_g),$$

and

$$P_{g\ ber} = 1/2 \cdot \rho (U/A_g)^2.$$ 

The oral constriction is modelled as a cylindrical tube of cross-sectional area $A_c$ and length $l_c$. The static pressure drop is then defined by:

$$P_{c\ stat} = P_{c\ lam} + P_{c\ ber} + P_{c\ tur},$$

where

$$P_{c\ lam} = 8 \pi l_c \cdot (U/A_c^2),$$

$$P_{c\ ber} = 1/2 \cdot \rho (U/A_c)^2,$$

and

$$P_{c\ tur} = K \cdot l_c \cdot (U^2/4) / (A_c^2 \cdot A_c^3/8),$$

with

$$K = (\rho/4) \cdot (\pi^5/2 \cdot v/200)^{1/4}.$$ 

From these equations, given $P_{sgl}=P_{g\ stat}+P_{c\ stat}$, as well as $A_g$ and $A_c$, we can compute the flow $U$ and the different resistances, and finally derive the dynamic resistances as differential resistances.


2.2 Matching of measured spectra with simulated transfer functions

To improve the idealized modelling proposed earlier (Fant, Lin, & Badin, 1988), we have made an attempt to replicate the spectra shown in Fig. 2 from more realistic vocal tract area functions, taking also into account glottal and subglottal impedances.

We exemplify here the case of $[J]$, by far the most complex, due to the significant coupling between the front and back cavities.
2.2.1 Aerodynamic conditions
From the mask measurements, we know the aerodynamic conditions at the constriction but not at the glottis. We have thus chosen $A_c=0.3 \text{ cm}^2$, $U=350 \text{ cm}^3/\text{sec}$, and $\Delta p=0.8 \text{ cm H}_2\text{O}$ as representative values for the constriction; this choice leaves, however, the total aerodynamic conditions not completely specified. Fig. 10 shows the covariation of $P_{sgl}$ and $A_g$ that keeps IOP and flow constant; the corresponding glottal resistance, which is a good image of the subglottal coupling, is also shown. The following simulations were achieved with the arbitrarily chosen conditions $A_g=0.10 \text{ cm}^2$ and $P_{sgl}=8 \text{ cm H}_2\text{O}$ that correspond to a rather closed glottis and a weakly coupled subglottal system. They might not be realistic for unvoiced fricatives, but have been selected as a reference.

2.2.2 Determination of the area function
The first four important resonances of the [j] are $F_1=430 \text{ Hz}$, $F_2=1750 \text{ Hz}$, $F_3=2680 \text{ Hz}$, and $F_4=3200 \text{ Hz}$, and the first three antiresonances are $Z_1=1440 \text{ Hz}$, $Z_2=2270 \text{ Hz}$, and $Z_3=2980$. $F_1$ is necessarily the *Helmholtz resonance* between the back cavity and the constriction tube. $F_4$ is associated with the front cavity quarter wavelength resonance (slightly tuned by the constriction tube). $F_2$ and $F_3$ are the back cavity first two resonances above $F_1$, and $Z_1$ and $Z_3$ their bound zeros. $Z_2$ is the free zero corresponding to the resonance of the constriction inductance in parallel with the capacitance of the air volume behind the source.

![Graph showing subglottal pressure and glottis area covariations for constant IOP and flow. (A_c=0.3 cm^2, Δp= 0.8cm H_2O, U=350 cm^3/sec.)](image)
Fig. 11. Measured spectra, simulated transfer functions and corresponding area functions. (Weak subglottal coupling: $P_{sgl}=8.0 \text{ cm H}_2\text{O}, A_g=0.10 \text{ cm}^2$.)
From these observations, we could establish by trial and error a simplified area function. The top of Fig. 11 shows the rather good correspondence between the poles and zeros mentioned above; the second resonance of the front cavity (split by a bound pole/zero pair) has it cognates in the natural spectrum also (even though slightly too high in frequency). This proves that the proposed area function catches the most important acoustic features. The main overall spectral difference could be ascribed to a specific noise source spectrum falling off at 12 dB/oct above 2 kHz. However, our assumptions concerning the essential cavity structures of [ ] might not be representative.

For [s] and [f] (see Fig. 11), the fits were easy to achieve; the coupling with the back cavity seems weak, and the source spectrum could be assumed flat.

2.3 Influence of different parameters

2.3.1 Influence of source location: obstacle effect

It is well known from experiments, e.g., Shadle (1985), that an obstacle in a tube creates more intense noise than what is generated by an upstream constriction. To what extent is this predictable from the location of the source at a constant source pressure level? Can linear source-filter theory and essential VT cavity features explain not only characteristic spectral shapes but also overall spectral levels?

We have focused the analysis on essential cavity structures omitting the effect of cavities posterior to the constriction by means of an acoustic short circuit.

Our reference model has a front cavity of 2 cm length and a constriction of 1 cm length. At the left of Fig. 12 we have simulated a dental source location 1.5 cm anterior to the constriction outlet. The transfer function is remarkably insensitive to variations in constriction area.

On the other hand, with the source located at the anterior end of the constriction, the transfer function overall level varies proportionally to the constriction area (right part of Fig. 12). This can be thought of as a source impedance effect. Could this have influenced derivations of source dependency on \( A_c \)? To us it seems logical to treat it as a filter rather than as a source-property. Moreover, for a realistic value of the constriction area, \( A_c = 0.125 \, \text{cm}^2 \), the peak level of the main formant is about 20 dB below that of the dental source. This finding could in part account for the well known greater efficiency of an obstacle source. The obstacle effect, i.e., the enhancement of the overall radiated sound pressure level in the case of a dental source, is strongly supported by our simulations. It should anyhow be added that this effect is related to the ratio between the front cavity area \( A_{FC} \) and the constriction area \( A_c \): our simulations have shown that it is much weaker for \( A_{FC} = 1 \, \text{cm}^2 \) and \( A_c = 0.3 \, \text{cm}^2 \) than for \( A_{FC} = 2 \, \text{cm}^2 \) and \( A_c = 0.125 \, \text{cm}^2 \).

The lower middle part of Fig. 12 shows the effect of intermediate source locations. As the location is moved towards the teeth a spectral zero moves towards lower frequencies thus enhancing the main spectral peak. This zero can be thought of as a resonance of the constriction inductance in parallel with the capacitance of the air volume behind the source.

2.3.2 Influence of glottis opening

We had arbitrarily chosen \( A_g \) and \( P_{sgl} \) for the simulation of [ ]. In the measured spectrum, we can distinguish clearly two small extra resonances near 1 kHz and 2.2 kHz; these resonances are very likely due to subglottal coupling. Fig. 13 shows that when the glottis area increases (the subglottal pressure covarying as in Fig. 10), extra resonances appear, due to subglottal coupling. We could thus simulate these extra peaks by tuning \( A_g \) to 0.25 cm\(^2\) and \( P_{sgl} \) to 2 cm H\(_2\)O, keeping the same aerodynamic conditions at the constriction (see Fig. 14).
Dental source: transfer functions.
\( A_g = 1 \text{cm}^2; A_c \) from 1. to 0.008 \text{cm}^2

Front cavity resonance levels.
\( P_{\text{sc}} = 8 \text{cmH}_2\text{O}; A_g = 1 \text{cm}^2 \)

Constriction source: transfer functions.
\( A_g = 1 \text{cm}^2; A_c \) from 1. to 0.008 \text{cm}^2

Dental source:
area function.

Shift of source location: transfer functions.
\( A_g = 1 \text{cm}^2; A_c = 0.125 \text{cm}^2; X_{\text{src}} = 0, 0.5, 1, 2 \text{ cm} \)

Constriction source:
area function.

---

**Fig. 12** Effect of constriction area variations on transfer functions for dental and constriction sources.
Fig. 13. Influence of glottis area on transfer functions for [ɪ].
(A and P_{sgl} covarying as in Fig. 10 - Curve spacing = 20 dB.)

Fig. 14. Measured spectrum and simulated transfer function for [ɪ].
(Strong subglottal coupling: P_{sgl} = 2 cm H₂O, A_g = 0.25 cm².)

From this, we see the importance of the coupling between the front and back cavities for the fricative [ɪ]; the double peak around 3 kHz can be explained only in this way, and this result is reconfirmed by a number of bound pole/zero pairs that appear in the measured spectra.

2.3.3 Influence of constriction area and length
Fig. 15 shows the effects of constriction area variations upon the transfer function of [ɪ]; they differ from those adopted in the simulations described in Section 2.3.1 by the back cavities being taken into account. Due to a finite coupling, the frequency of the front cavity main res-
Resonance is shifted, but its amplitude remains fairly constant. We also note that, when the constriction area increases, bound pole/zero pairs appear, as a consequence of a greater front/back cavity coupling. Nevertheless, this reconfirms our explanations of the obstacle effect.

![Graph showing influence of constriction area on transfer functions.](image)

**Fig. 15.** Influence of the constriction area on transfer functions for $|\tilde{t}|$. ($A_c$ increases linearly from 0.1 to 0.6 cm$^2$.)

As can be seen in Fig. 16, the main influence of the constriction length is to shift of the frequency of the front cavity resonance, which is tuned by the constriction tube. We also notice the increase of front/back cavity coupling when the length decreases.

![Graph showing influence of constriction length on transfer functions.](image)

**Fig. 16.** Influence of the constriction length on transfer functions for $|\tilde{t}|$. ($l_c$ increases linearly from 0.3 to 1.2 cm; curve spacing = 20 dB.)
2.4 Conclusions
Our modelling of [s] and [f] has been quite successful: we have obtained a good fit between the measured spectra and the computed transfer functions. However, the overall fit was not completely satisfactory for the [j] sound: our area function modelling might not be representative, and we still lack realistic data on cavity configuration and source location in that case. We have shown the importance of the constriction dimensions for the front/back coupling typical of [j]. We have also demonstrated that the well known obstacle effect could be explained in terms of filter theory.

In order to advance our knowledge, simultaneous X-ray pictures and aerodynamic measurements, as well as measurements of the glottal area and flow, are planned.

Acknowledgments
This work, realized when the author was a guest scientist at KTH, has been supported by the French Centre National de la Recherche Scientifique (CNRS) and the Swedish Board for Technical Development (STUF), in the frame of an international cooperation program.

The author wishes to express his deep and warm gratitude to Gunnar Fant, who has been a very devoted and encouraging guide, and has also kindly and efficiently reviewed an earlier version of this paper.

The author is also very much indebted to many of his colleagues and friends at KTH, for their help with recordings, computer programs and management, edition ... They will find here his most sincere thanks.

References
LIST of SYMBOLS

A        Duct diameter
A_c      Cross-sectional area, or minimum constriction area
c        Sound velocity in air
d        "Characteristic dimension of the constriction"
Δp       Pseudo-static pressure drop across the constriction (Identical to IOP)
f        Frequency
K(S_t)   Parameter dependant on the Strouhal number
l_c      Constriction length
μ        Viscosity of air
v        Kinematic coefficient of viscosity (=μ/r)
p        "IOP" exponent (relation P_r ~ (Δp)^P)
P_r      Radiated Sound Pressure (RMS average of the microphone pressure)
P_s      Noise source pressure (RMS value)
P_sgl    Subglottal pressure
q        "Area" exponent (relation P_r ~ (A_c)^q)
r        "velocity" exponent (relation P_r ~ (V_c)^r)
ρ        Air density
SPL      Sound Pressure Level (=20log_{10}(P_r))
S_t      Strouhal number (fd/V_c)
V_c      Volume velocity in the vocal tract
V_p      Particle velocity in the constriction (=U/A_c)
W_{Δf}  Power radiated in a duct by a noise source (in 1/3 octave bands)