FRICATIVE PRODUCTION MODELLING: 
AERODYNAMIC and ACOUSTIC DATA.

Pierre BADIN* Gunnar FANT**

*Institut de la Communication Parlée (UA CNRS), INPG - Univ. Stendhal
INPG : 46, av. F. Viallet, F-38031 Grenoble Cedex, FRANCE
**Department of Speech Communication and Music Acoustics, KTH
Box 70014, S-100 44 Stockholm, SWEDEN

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 pseudo-static 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 transfer functions 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 aerodynamic and acoustic phenomena involved in the production of fricative consonants are far from being completely understood. In order to improve our knowledge in that field and our fricative production models, we have attempted to obtain new experimental data on a human subject. In the first part of this paper, we deal with characteristics related to vocal tract aerodynamics and in the second part, we report simulation studies more related to acoustics.

1. AERODYNAMICS

The overall level and spectral characteristics of the radiated sound pressure primarily depend on the aerodynamic state of the vocal tract, namely the pseudo-static pressure drop across the main oral constriction Δp and its cross-sectional area Ac. We report here data obtained for human sustained and dynamic fricatives.

1.1 EXPERIMENTS and RESULTS

Experimental setting - Data processing

We replicated and extended the experiments reported in [1] on human fricatives. We used a circumferentially vented pneumotachograph mask [2], which measures oral flow U, and Intra Oral Pressure (IOP) Δp through a polyethylene tube inserted in the mouth. Due to distortion brought by the mask to the radiated sound, we split the measurements into two series : (1) simultaneous recording of radiated sound pressure, IOP and flow ; (2) sound pressure and IOP only (without mask). For both sustained and dynamic fricatives, the corpus was limited to three voiceless fricative consonants [f], [s] and [V], produced by a French subject, PB.

In one set of experiments, the subject was instructed to sustain the syllables [a], [as] and [aC], keeping the fricative segment for at least 500 ms, at different levels from the lowest effort to the highest possible. The absolute SPL was computed from the microphone sound pressure RMS average SP, and the IOP and the flow from a simple averaging over the total segment lengths. The spectra of the sound pressure recorded for each item in the experiment without mask were estimated from the Long Time Average Spectrum method (see Fig.4).

For the study of fricative dynamics, a number of [CataC] logatoms were recorded. The consonant C was either [f], [s] or [V]. For each logatom, the voiceless parts of [Ca] and [aC] were edited, in order to measure the SPL, Δp and Ac data of [f], [s] and [V]. For each logatom, the voiceless parts of [Ca] and [aC] were edited, in order to measure the SPL of the noise component only, excluding any possible voiced component in the transitions. The SPL was evaluated from RMS data of successive 5ms windows ; SPL, IOP and flow signals were low-pass filtered at 80 Hz, and the constriction area computed according to the orifice equation. Finally the signals were undersampled at 1 kHz, to produce, for each fricative, sets of measured SPL, Δp and Ac data of 100-300 points.

Aerodynamic parameters

From the IOP and the flow, we evaluated the minimum constriction area from the orifice equation [31] :

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

(1)

For [V], we found minimum constriction areas varying between 0.1 and 0.4 cm². The subject seemed 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. For [s] and [f] the constriction areas were approximately independent of the IOP, around 0.1 cm². In general, all our results check fairly well with [1], if we notice that they used a slightly different orifice equation, providing 40% lower areas. Our data for sustained and dynamic fricatives are also consistent with each other.

Overall level of radiated sound pressure

We determined from the data on sustained fricatives (mask series) the "IOP" and "Area" exponents, i.e. the parameters p and q in the relation :

\[ SP \sim (\Delta p)^p (A_c)^q \]  

(2)
SPL = 20 log(SP) = G₀ + p.20 log(Δp)+q.20 log(Ac),

by applying multiple linear regression to the three fricative
classes, where the dependent variable was SPL. From the same
data we also determined the "IOP" exponents assuming no Ac
dependency. From the equivalent set for the series without mask
the "IOP" exponents were also computed. These results are
given in Table 1.

For the dynamic fricatives, multiple regression was applied to each set of (SPL, Δp, Ac) data, in order to determine
the "IOP" and "Area" exponents for each individual consonant.
These exponents were also estimated globally for each fricative
class; they are given, as well as those derived from the IOP data
only, neglecting the influence of Ac, in Table 2.

Spectrum of the radiated sound pressure

We have found that the high frequency part of the
sustained fricative spectrum increases with overall SP more
rapidly than the low frequency part. In order to measure more
quantitatively this change, we have computed the SP in ten
frequency bands of approximatively 1 kHz widths, and
estimated for each band the "flow" exponents (Table 1). The
fricative [J] exhibits the greatest spectrum tilt change, whereas
[J] shows the smallest, and less regular changes.

Fig. 2 reconfirms these spectral tilt changes related to
SPL for dynamic fricatives: high frequencies tend to
appear shortly after the low frequencies onset during the transition into
the fricative, whereas they disappear before the low frequencies
during the transition out of the fricative.

This phenomenon presumably reflects similar variations
in the source spectrum.

1.3 DISCUSSION

The assumption of a direct proportionality between the
overall source and the overall radiated sound pressure holds
only for source signals having a uniform spectral dependency on
Δp and Ac, 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, since
all the experimental data in the literature, as well as ours, have
been obtained from measurements on the radiated sound
pressure, and not on the source itself, we will discuss the comparisons on the radiated sound pressure basis.

Simulations have shown that, in our experimental
conditions, the major part of the IOP is due to the
BERNOULLI pressure drop: we are thus entitled to assume in the
following that Δp-Vc², neglecting the loss terms, and we
can therefore convert the "IOP" exponents to "flow" exponents.

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

The comparison between Table 1 and 2 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 (cf for instance maxima and minima in Table 2). To get a better evaluation of these coefficients, we have
resynthesized, for each spoken fricative, the SPL dynamic
curve as a function of Δp and Ac, from equation (3), using for
g₁₀, p and q three different sets of values obtained by multiple
regression in different conditions: (1) SPL against Δp and Ac
for each individual fricative, (2) SPL against Δp and Ac,
globally for each fricative class, and (3) SPL against Δp only,
for each fricative class (see Fig. 3). A visual inspection of such
curves for all fricatives shows that, if we suppose removed the
residual oscillations of the measured curve (due to short term
noise fluctuation), the maximum departure from the measured
curve is within 5-6 dB. Moreover, we found that in most cases for [J] and [s], the fit with an Ac dependency is better than the
one without; this reconfirms this dependency.

The "flow" exponents derived from our data, as well as
the corresponding data from [4] for obstacle and non obstacle
configurations are given in Table 1. The agreement with [4] is
fairly good, and we note that our [J] and [s] have "flow" exponents close to the value 2.5 obtained by [4] for her obstacle
configuration with a 3.2 cm long front cavity.

From the comparison, it comes also out that our [J] behaves rather as an obstacle configuration than a non-obstacle one. Our subject reported that during the production of [J], 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 "flow" exponent for [J].

1.4 CONCLUSIONS

The previous analysis shows that, in all cases, the "flow"
exponents that we have derived lie between 2 and 3, for [J] and
[s], with "Area" exponents between 0 and 0.5. These values
correspond to a situation intermediate between the relations
SPL-Vc² proposed by [5] and SPL-Vc³/3Ac proposed by [6].
We suggest thus to use p=1.3 and q=0.3 for [J] and [s], and
p=0.8 and q=0.2 for [f] which behaves in a fairly different way.

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 tilt changes.
We should also substitute the IOP probe for one inserted
through the nasal cavity, to treat also the most retracted
consonants. It will be also of great interest to apply these
equations to the source itself in a synthesis scheme, using the
proper spectrum dependent corrections.

2. ACOUSTICS

This section deals with the acoustic modelling of the
fricative production. We assume aerodynamic pressure losses at the
glottis competing with the constriction pressure drop (see for instance [6]), and we include frictional losses as the dominating
resistance at very small openings. The acoustic modelling itself
uses a representation of the vocal tract by electric quadrupoles
which leads to vocal tract frequency transfer functions (cf [7]).

2.1 MATCHING OF MEASURED SPECTRA WITH
SIMULATED TRANSFER FUNCTIONS

To improve the idealized modelling proposed by [8], we
have attempted to replicate some measured spectra from more
realistic vocal tract area functions, taking into account back
cavity, glottal and subglottal impedances.

We exemplify first the case of [J], by far the most
complex, due to the significant coupling between the front and back cavities. In the measured spectrum (Fig.4a), the first four important resonances are F1=430Hz, F2=1750Hz, F3=2680Hz and F4=3200Hz, and the first three antiresonances are Z1=1440Hz, Z2=2270Hz and Z3=2980. F1 is necessarily the Helmholtz resonance between the back cavity and the constriction tube, whereas F4 is associated with the front cavity quarter wavelength resonance. F2 and F3 are the back cavity first two resonances above F1, and Z1 and Z3 their bound zeroes. Z2 is the free zero corresponding to the resonance of the constriction inductance in parallel with the capacitance of the air volume behind the source. From these observations, we could establish by trial and error a simplified area function. Fig.4 (a, c, d) shows the rather good fit between synthesized and measured spectra: 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 [j] might not be representative. For [s] and [f], the fits were easy to achieve; the coupling with the back cavity seems very weak, and the source spectrum could be assumed flat.

2.2 INFLUENCE OF DIFFERENT PARAMETERS

Source location and constriction area: obstacle effect

It is well known from experiments, e.g. [4], 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 vocal tract 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 acoustical 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.5 we have simulated a dental source location 1.5 cm anterior to the constriction outlet. The transfer function is remarkably insensitive to variations in construction 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. 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\(_c\)=0.125cm\(^2\), 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\(_F\) and the constriction area A\(_c\): our simulations have shown that it is much weaker for A\(_F\)=1cm\(^2\) and A\(_c\)=0.3cm\(^2\) than for A\(_F\)=2cm\(^2\) and A\(_c\)=0.125cm\(^2\).

Glottis opening and subglottal system

We had arbitrarily chosen A\(_g\)=0.13cm\(^2\), and P\(_{\text{subgl}}\)=1.0cmH\(_2\)O for the simulation of [f], to fit the aerodynamic values observed from the mask measurements: A\(_c\)=0.3cm\(^2\), U=400cm\(^2\)/sec and D\(_p\)=0.8cmH\(_2\)O. On the measured spectrum we can distinguish clearly two small extra resonances near 1. and 2.2 kHz; those are very likely due to subglottal coupling. We could thus simulate these extra peaks by tuning A\(_g\) to 0.3cm\(^2\) and P\(_{\text{subgl}}\) to 2.4cmH\(_2\)O, keeping the same aerodynamic conditions at the constriction (see Fig.4b).

From this, we see the importance of the coupling between the front and back cavities for the fricative [j]: 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 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. Simultaneous X-ray pictures and aerodynamic measurements are under way to fulfill this need.

ACKNOWLEDGEMENTS

This work, realized when Pierre BADIN was a guest scientist at KTH, has been supported by grants from the Swedish Board of Technical Development (STUF), the Swedish Telecom, and the French Centre National de la Recherche Scientifique.

REFERENCES


<table>
<thead>
<tr>
<th>Source Location</th>
<th>&quot;Area&quot; Exponents for Human Sustained Fricatives</th>
</tr>
</thead>
<tbody>
<tr>
<td>[j]</td>
<td>1.32 1.24 0.84 1.27 1.4 1.2</td>
</tr>
<tr>
<td>[s]</td>
<td>1.26 1.38 0.72 1.42 1.3 1.3</td>
</tr>
<tr>
<td>[f]</td>
<td>1.04 0.97 0.81 1.31 -</td>
</tr>
<tr>
<td>(1) &quot;IOF&quot; exponents for PB (Mask series)</td>
<td>0.65 (0.09/0.68)</td>
</tr>
<tr>
<td>(2) &quot;IOF&quot; exponents for PB (No-Mask series)</td>
<td>0.65 (0.09/0.68)</td>
</tr>
<tr>
<td>(3) &quot;IOF&quot; exponent for TH (Pb)</td>
<td>0.65 (0.09/0.68)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Source Location</th>
<th>&quot;Area&quot; Exponents for Human Dynamic Fricatives</th>
</tr>
</thead>
<tbody>
<tr>
<td>[j]</td>
<td>1.24 0.98/1.66 0.42 (-0.09/0.68) 1.26</td>
</tr>
<tr>
<td>[s]</td>
<td>1.42 0.81/2.09 0.23 (-0.19/0.82) 1.31</td>
</tr>
<tr>
<td>[f]</td>
<td>1.77 0.62/1.68 0.27 (-0.20/0.99) 0.75</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Source Location</th>
<th>&quot;Area&quot; Exponents for Human Sustained Fricatives</th>
</tr>
</thead>
<tbody>
<tr>
<td>[j]</td>
<td>1.24 0.98/1.66 0.42 (-0.09/0.68) 1.26</td>
</tr>
<tr>
<td>[s]</td>
<td>1.42 0.81/2.09 0.23 (-0.19/0.82) 1.31</td>
</tr>
<tr>
<td>[f]</td>
<td>1.77 0.62/1.68 0.27 (-0.20/0.99) 0.75</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Source Location</th>
<th>&quot;Area&quot; Exponents for Human Dynamic Fricatives</th>
</tr>
</thead>
<tbody>
<tr>
<td>[j]</td>
<td>1.24 0.98/1.66 0.42 (-0.09/0.68) 1.26</td>
</tr>
<tr>
<td>[s]</td>
<td>1.42 0.81/2.09 0.23 (-0.19/0.82) 1.31</td>
</tr>
<tr>
<td>[f]</td>
<td>1.77 0.62/1.68 0.27 (-0.20/0.99) 0.75</td>
</tr>
</tbody>
</table>

The exponents are estimated over the whole mask data set; in parenthesis are given the minimum and maximum values of the individual estimations.
Fig. 1: "Flow" exponents for human fricatives and mechanical models.

\[ \text{[f]} = "+", \text{[s]} = "o", \text{[t]} = "-" \]

SHADLE's obstacle case with front cavity length 3.2 cm = "x".

Fig. 3: Comparison between measured and calculated SPL for the dynamic fricative [f].

- (1) SPL measured on an individual fricative
  \( G_0, p, q \) derived from:
- (2) individual fricative (\( G_0=53.9, p=1.41, q=0.47 \))
- (3) whole [f] class (\( G_0=53.1, p=1.34, q=0.42 \))
- (4) whole [f] class (no \( A_c \) depend. assumed) (\( G_0=47.0, p=1.26, q=0.00 \))

Fig. 2: Spectrogram and spectral sections for [a].

(a) spectrogram; (b) spectral sections every 2 ms (from 4 to 34 ms). In both cases, the analysis bandwidth is 250 Hz.

Fig. 4: Measured spectra, simulated transfer functions and corresponding area functions.

(a) [f], (c) [s], (d) [t] (Weak subglottal coupling: \( P_g=8.0 \text{cmH}_2\text{O}, A_c=0.13 \text{cm}^2 \)); (b) [f] (Strong subglottal coupling: \( P_g=2.6 \text{cmH}_2\text{O}, A_c=0.3 \text{cm}^2 \)).

Fig. 5: Effect of constriction area variations on transfer functions for dental and constriction sources.

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

Front cavity resonance levels.
\( P_g=0.6 \text{cmH}_2\text{O}; A_c=1 \text{cm}^2 \)

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