Some geometric and acoustic properties of the lip horn

Pierre Badin,* Kunitoshi Motoki,** Nobuhiro Miki,***
Diane Ritterhaus,**** and Med-Tahar Lallouache*

*Institut de la Communication Parlée, URA CNRS No.368, INPG—Université Stendhal,
Grenoble, France
**Faculty of Engineering, Hokkai-Gakuen University,
Sapporo, Japan
***Faculty of Engineering, Hokkaido University,
Sapporo, Japan
****Technical University,
Dresden, Germany

(Received 11 May 1993)

In speech production research, a good knowledge of the articulatorily-acoustic relationship is a key factor. As lips are one of the major vocal tract acoustic control points, it is of utmost importance to gain insights into the relationship between their geometric and acoustic characteristics. Therefore, video recordings and plaster replicas of the lips of a subject were made for a set of sustained vowel and fricative articulations. Video images of the replicas were also recorded. From these recordings, geometric labial horn parameters, such as labial horn depth and intralabial area, were computed. The acoustic radiation impedance of the replicas, at the coronal plane located at the lip corner, was measured by a two-microphone method, and the length and area of a single uniform cylindrical tube of identical impedance were determined through an optimization algorithm, leading to an acoustic equivalence of the lip horn. Various correlations between the geometric parameters have been confirmed. It has also been shown that the acoustic equivalent intralabial area linearly depends on the geometric area, whereas the equivalent length is roughly independent of the geometric length. It was concluded that lip corner position is the main determinant of the vocal tract length on the lip side.

Keywords: Vocal tract acoustics, Lip horn geometry, Lip radiation impedance

PACS number: 43. 70. Bk, 43. 72. Ct

1. INTRODUCTION

In the domain of speech production research, a good knowledge of the articulatorily-acoustic relationship is a key factor. Particularly, it has been established since long that the lips are, together with the oral constriction, one of the major acoustic control points of the vocal tract (cf. *e.g.* the three-parameter models proposed in Refs. 1–3), and evidence shown in Ref. 4)). It is thus necessary, in the frame of articulatorily modeling and vocal tract acoustics, to gain insights into the geometric and acoustic characteristics of the lips and into their relationships. Moreover, a representation of the complexity of the lip horn by means of a limited set of acoustic parameters would be of high interest for articulatory synthesis. Lip geometry has been given a constant interest by Abry and colleagues at the Institut de la Communication Parlée in Grenoble.5,6) Lip radiation characteristics have been studied in Refs. 2, 7), based both on experimental measurements on dummy heads and on adaptations of theoretical work in acoustics.8) However, works aimed at establishing detailed relationships between geometric and acoustic characteristics of lips have not been attested in the literature, except for that of
Motoki and colleagues at the Research Institute for Electronic Science in Sapporo. The study presented in this paper is an extension of the work of the latter on Japanese vowels, to French vowels and fricative consonants, with more emphasis on geometric measurements carried out on a real subject.

Because of the difficulty of directly measuring radiation impedance on the subject, plaster replicas of lips were obtained; the geometry of these replicas was then determined, as well as radiation impedance. Lip horn geometry for the sustained articulation of the same set of sounds produced by the subject was also analyzed, in order to verify the reliability of the replicas and to elaborate on knowledge of the range of variation of different parameters.

Experimental measurement methods of labial geometry and radiation impedance are described in the first part of this paper. In the second part, the relationships between these different parameters are investigated and discussed.

2. EXPERIMENTAL SETUPS AND METHODS

Due to the complexity of the different operations involved in the measurement procedures, the experiment was carried out on a limited number of productions. The sustained configurations under investigation were limited to some French vowels and fricative consonants uttered by a single French male subject. It is worth noting that the same subject had already been involved in a number of recordings made for the same productions (radiated sound pressure, pressure drop and flow at the oral constriction, electropalatography, teleradiography, video lip recordings from face and profile, etc.) in the frame of a European project,\(^{12,14}\) and that the present study constitutes a complementary approach in our search for understanding speech production phenomena.

2.1 Geometric Characterization of the Lips

The traditional approach in vocal tract acoustics is to consider the vocal tract as a non-uniform tube, approximately closed at the glottis and open at the lips. Acoustic waves traveling in this tube are partially radiated at the lips, and possibly through the vocal tract walls. The problem of characterizing lip geometry is related to the question: *Where does the vocal tract end?* Indeed, even though it is possible to define the limits of the vocal tract along its midline from the glottis to the lip corner—since the intersection between a plane perpendicular to the midline and the vocal tract walls is a closed contour—, it is however impossible to do so for the lip horn, *i.e.* the portion of the vocal tract anterior to the lip corner. It is important to recall that the way vocal tract geometry is described is strongly related to the acoustic simulation method adopted. A finite element description of the outline of the vocal tract as a two-dimensional surface can be envisaged, but would lead to rather complex simulations. Most vocal tract acoustic models are thus based on the hypothesis of plane waves propagating along the midline of the tract (see e.g. references in Ref. 15)). Under such an assumption, vocal tract geometry can be reduced to a cascade of uniform cylindrical tubes, whose dimensions define the area function, that may be simulated by electrical quadrupoles. The problem of mapping the midsagittal function to the area function has already been well addressed,\(^{16,17}\) but the specific problem of the lips has not been accurately solved yet.

The shape of the lip horn can be characterized, from front and profile pictures of the subject's face, by a set of eight parameters that were proposed by Abry and Boë\(^\circ\) (see Fig. 1):

- from front: lip height (B), lip width (A), intralabial lip area (S).

\(\text{Fig. 1 Lip parameters (from Lallouache, 1990). (a) profile view; (b) front view.}\)

\(A: \text{lip width, } B: \text{lip height, } S: \text{intralabial lip area, } P_1: \text{upper lip protrusion, } P_2: \text{lower lip protrusion, } C: \text{lip corner protrusion, } L: \text{lip horn depth, } D: \text{lip horn opening.}\)
P. BADIN et al.: GEOMETRIC AND ACOUSTIC PROPERTIES OF LIPS

- from profile: upper \((P_u)\), lower \((P_l)\) and corner \((C)\) lip protrusions, defined as the distances of these points to a vertical reference line (drawn on the ruler attached to the goggles worn by the subject, see Fig. 1); lip horn depth \((L)\) and opening \((D)\).

The parameters that are most likely acoustically relevant are lip area, and horn depth and opening. It is possible to automatically extract these parameters from video recordings using the system developed at ICP by Lallouache and colleagues.\(^{5,19}\) The measurement method, that involves two steps, namely video recordings from face and profile, and contour detection, will be described later.

2.1.1 Plaster replicas

Because of the impossibility of measuring radiation impedance directly on the subject, both the acoustic and the geometric measurements were made on plaster replicas of the lips of the subject. The material used in the present study thus consists primarily of lip replicas, and of front and profile video recordings of both the replicas and the subject himself.

The method adopted to obtain the plaster replicas is similar to techniques used in orthodontia for hard palate and teeth replica fabrication, that had been successfully utilized in a similar study.\(^{11}\) In order to avoid deformation of the lips, which are made up of soft tissues, the “negative” cast was obtained using an alginate paste made soft enough by adding an extra proportion of water. The subject, after having sucked in a sufficient quantity of paste into the anterior part of his mouth, adopted the desired articulatory position, and then covered, as much as possible, the outer part of his lips with the leftover alginate material. After a gelatinization time of about three minutes, the removal of the “negative” cast was relatively easy due to the rubber-like texture of the material, and then a “positive” replica was made out of plaster. This procedure allows obtaining very good quality replicas, faithfully respecting the shape of the entire lip horn (upper and lower lips), the teeth, and in some cases the constriction of front fricative consonants.

2.1.2 Video recordings

In order to obtain the best accuracy and reliability in the determination of the lip parameters, Lallouache and colleagues\(^{5,19}\) have developed a special setup and protocol for video recordings of a subject’s face (see Fig. 2).

The subject was seated in a soundproof room at an equal distance from two cameras (positioned for full-face and profile recording), with the head position fixed by a helmet that was mounted on the chair in such a way as to ensure its stability. The subject wore a pair of dark goggles which were held in a stable position by means of a band fixed around the back of the head. A ruler was attached to the goggles and provided a reference point for small head movements that occurred during speech production, whereas the goggles protected the subject from the lighting. The subject’s lips were painted blue (being careful enough not to color the teeth) so as to allow the lip contours to be extracted from the image without any distortion. This was done by means of an analog real-time preprocessing (Universal Chroma-key, SONY CRK 2000P) which allowed the extraction of a single color only. The chroma-key then substituted the blue color with a signal of zero luminance. In order to avoid any ambiguity between the gray level of the lips and that of the cavity between the lips, the level of black recorded by the camera was lightened slightly. After processing by the chroma-key, the lips thus appeared black and in 2D. The processed signals from the face and profile cameras were mixed and recorded on a UOMATIC-SP video recorder, using the PAL standard, at a field frequency of 50 Hz. An audio recording was carried out simultaneously. A microphone, positioned at about 20 cm from the subject’s lips, was connected directly to channel 1 of the video recorder, and to channel 2 through a “synchronizer” that allowed the synchronized recording of a LED
flash on the video channel and a 3 kHz beep on the audio channel. The subject triggered this synchronization signal at the beginning of the production of each item of the corpus by means of a push-button. The LED was positioned behind and to one side of the subject.

The recording protocol for the replicas was slightly different from that of the subject. The plaster replicas, where the lips had been carefully painted blue, were hung onto a vertical screen installed at the subject's previous place. The video setup was the same, but the images, instead of being recorded into the video recorder, were directly digitized through the image processing board and stored on the PC hard disk. In order to allow exact normalization, a reference ruler image was also recorded.

2.1.3 Determination of the lip parameters

Each video field was extracted and digitized over 512 x 512 alternating pixels coded with 256 gray levels. Three contours were then determined from each field: two closed contours, the internal and external contours of the lips as seen from full-face, and one open contour as seen from profile. From these contours, the parameters $A$, $B$, $S$, $C$, $P_0$, $P_1$, $L$, and $D$, were automatically extracted by a computer program, and stored for further processing or reference. Intralabial lip area $S$ was obtained by a pixel counting procedure. Lip height $B$ was calculated at the point of center of gravity, rather than at the center of the horizontal dimension because of asymmetries in the inter-lip cavity shape, with the help of a quadratic interpolation of the upper and lower lip curves. Parameters $A$, $C$, $P_0$, $P_1$, $L$, and $D$, were determined from simple geometric properties such as tangency or extremum of a curve.

2.2 Acoustic Characterization of the Lips

To perform an acoustic simulation of the vocal tract, it is necessary to know, not only the area function, but also the vocal tract boundary conditions, especially the radiation impedance which is the acoustic load that takes into account the effects of the ambient field around the lips on the acoustic waves inside the tract. It is possible to derive analytically this impedance in some idealized cases, such as a small piston set in a spherical or plane baffle. The physical interpretation of the imaginary part of the complex radiation impedance is the end correction, i.e. the virtual lengthening of the last tube of the vocal tract, which influences the resonance frequencies, especially those of the front cavity for front consonants. The real part of the radiation impedance corresponds to the radiation losses, i.e. the energy radiated by the vocal tract.

2.2.1 Defining the object to be measured

It was arbitrarily decided to use the coronal plane containing the contact point of the lips (lip corner) as a reference plane for defining the lip radiation impedance; the radiation impedance thus represents the entire lip horn. Under the assumption of plane waves propagating inside the vocal tract, i.e. up to the reference plane (this does not imply that the waves are plane downstream this point), the radiation impedance is the acoustic impedance seen from the vocal tract in the direction of the lip horn, at the reference plane. The radiation impedance is thus the ratio between the complex acoustic pressure and the complex acoustic flow entering the lip horn. A modified version of the two-microphone method has been used to determine this impedance as a function of frequency.

2.2.2 Acoustic measurement principle

The theoretical principle consists in imposing a set of stationary waves in an acoustical tube with a uniform section. The tube is loaded at one end by the object whose impedance $Z_o(\omega)$ is to be measured, and excited by means of a sinusoidal source at the other end. Two microphone probes are placed inside the tube at locations of coordinates $x_1$ and $x_2$; the amplitude of the pressure signals and their relative phase are measured with the help of a frequency analyzer. It is assumed that in the center part of the uniform tube, there are standing plane waves: the two microphones are placed far enough from the ends of the tube such that they are situated in this region. The following formula can thus be used to describe the state of transmission:

$$Z_\omega(\omega) = \frac{\sinh(\gamma \cdot x_2) - H \cdot \sinh(\gamma \cdot x_1)}{H \cdot \cosh(\gamma \cdot x_1) - \cosh(\gamma \cdot x_2)} Z_0$$

where $Z_\omega(\omega)$ is the input impedance at coordinate $x = 0$ (i.e. the radiation impedance), $H$ is the complex transfer function between the pressures measured at points $x_1$ and $x_2$ (i.e. $H = P(x_1)/P(x_2)$), $\gamma$ is the propagation constant, and $Z_0$ the complex characteristic impedance. Classical viscous and thermal losses are assumed in the uniform tube, and the air density and sound velocity are corrected for temperature.
In the actual setup, the pressures were measured at points $x_1$ and $x_2$ using only one microphone probe that was successively moved to each of these points, instead of using two fixed microphone probes: problems related to linearity and calibration of the microphones were thus eliminated. Finally, the only restriction for microphone quality was fidelity.

2.2.3 Experimental setup

The experimental setup depicted in Fig. 3 is roughly that used by Motoki et al.\textsuperscript{113} The replica was fixed on a baffle of large dimensions. In order to ensure a better fit to the shape of the lips, the acoustical uniform plexiglass tube was given a pseudo-elliptical cross-section. The plexiglass tube was inserted into a cavity that was bored into the replica from the rear side. The depth of the cavity was adjusted, for each replica, in such a way that the end of the tube was approximately 1 mm behind the lip corner. This means that the teeth were removed for all the configurations where the teeth were located behind the lip corner; thus typically $[i, f, \theta]$ were the only configurations including teeth. The microphone probe was made of a small glass tube (1.8 mm inner diameter, 3.0 mm outer diameter, 500 mm length) of which one end had been deformed in order to fit a small ordinary necktie microphone. This probe was fixed to a transverse device which could be moved longitudinally, with an absolute positioning accuracy better than 0.01 mm, and thus allowed pressure measurement at any desired location inside the tube. The generator component of a frequency analyzer was used to generate a sinusoidal signal which fed the loudspeaker that maintained the standing waves in the tube. The signal of the microphone probe was preamplified and sent to the frequency analyzer, which measured amplitude and phase relative to the signal picked up by another fixed probe microphone that was inserted in the uniform tube, 100 mm from the replica-side end of the uniform tube. This choice of reference signal aimed at diminishing any undesired influence due to slight changes in experimental conditions, such as room temperature, during measurement. This reference was deemed better than the excitation signal since the phase lag from the loudspeaker to the replica, which is sensitive to a small change in room temperature, was excluded by use of this reference signal. The experimental setup was entirely controlled by a minicomputer that drove the frequency analyzer to produce a sequence of excitation signals (frequency and amplitude) and stored values of amplitude and phase measured by the analyzer for each frequency point (typical measurement errors for the analyzer were $\pm 0.03$ dB in amplitude and $\pm 0.05$ degrees in phase). Magnitude and phase of the impedance were then automatically calculated for each frequency point. One example of radiation impedance, represented as the amplitude and phase of the reflection coefficient of the acoustical flow in the reference plane $\mu(\omega) = [Z_0 - Z_L(\omega)]/[Z_0 + Z_L(\omega)]$, is given in Fig. 4 (thick lines).

2.2.4 Determining an acoustic equivalence of the radiation impedance

The acoustic characterization of the lip horn can be fully acquired knowing the radiation impedance seen from the coronal plane that includes the lip corner. However, it would be conceptually easier to represent this impedance, which acts as one boundary condition of the vocal tract, by an \textit{acoustic equivalence} defined as a cascade of uniform tubes having the same input impedance. Such an equivalence can be determined with the help of an optimization algorithm. Optimization consists in minimizing a distance (both in amplitude and phase, cf.,\textsuperscript{113} for more details) between the reflection coefficient experimentally measured and the reflection coefficient computed from the cascade of uniform tubes, over a number of frequency points. As experience has shown that the convergence of the optimization algorithm is uncertain for a cascade of several sections, we decided to search for a one-section equivalence. The advantage is that the
form measurements on replicas of the subject’s lips. Replicas were made for the following sustained vowels and fricative consonants: [i, ø, o, u, f, ð, j]. As lip shape for [s, ç] are rather similar to that of [i], no replicas were made for these consonants. Geometric and acoustic measurements obtained for the replicas are presented in Table 1.

In order to verify the reliability of the replicas and to elaborate on knowledge of the range of variation of different parameters, video recordings were made for the fricatives [f, ð, s, j, ç]. A stable fricative position was ensured by uttering the item immediately after an [a] vowel, for example [affiff] for all consonants except for [ç] that was preceded by vowel [i]. Table 2 gives average geometric measurements obtained for the stable central portion of these sustained configurations.

3.2 Statistical Analysis of the Data

Correlation matrices were computed for the different parameters, including the product A·B, for both replicas and the subject’s measurements (see Tables 3 and 4). The analysis of these results leads to a number of observations.

3.2.1 Geometric parameters

The most striking correlation is that of geometric intralabial area with the product of lip width A by lip height B, both for the replicas and the subject. This is a well established fact.20,21) The slopes of the regression lines of S as a function of A·B have been computed for the replicas data in Table 1, and for the subject’s data of which the averages are given in Table 2. These slopes are 0.84 for the replicas and 0.74 for the subject, with correlation coefficients of 0.977 and 0.995 respectively.

Another strong correlation is that of lip height B with lip horn opening D. It is verified that this correlation is clearly due to a relationship between the different phonemic classes, but does not hold very well within each class; thus it is not related to a physical phenomenon, but to geometric properties that characterize the different phonemes studied here. This result had also been established by Abry and Boë22) for another corpus.

It appears also that lip horn depth L is fairly well correlated with lip width A (see also Figs. 5 and 6), especially for the subject. As can be seen from Fig. 5, this correlation (0.940) is also related to a relationship between classes, but does not hold at all within a class. It simply reveals that [f], for in-

It appears also that lip horn depth L is fairly well correlated with lip width A (see also Figs. 5 and 6), especially for the subject. As can be seen from Fig. 5, this correlation (0.940) is also related to a relationship between classes, but does not hold at all within a class. It simply reveals that [f], for in-
Table 1 Measured geometric and acoustic parameters of the replicas.

<table>
<thead>
<tr>
<th>Phoneme</th>
<th>A (mm)</th>
<th>B (mm)</th>
<th>S (geom.) (mm²)</th>
<th>L (geom.) (mm)</th>
<th>D (mm)</th>
<th>S (acoust.) (mm²)</th>
<th>L (acoust.) (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>i</td>
<td>46.1</td>
<td>10.6</td>
<td>408.9</td>
<td>17.9</td>
<td>22.9</td>
<td>413.0</td>
<td>10.7</td>
</tr>
<tr>
<td>i*</td>
<td>49.7</td>
<td>12.5</td>
<td>485.6</td>
<td>17.5</td>
<td>24.8</td>
<td>555.0</td>
<td>11.0</td>
</tr>
<tr>
<td>ì</td>
<td>46.6</td>
<td>10.6</td>
<td>380.3</td>
<td>18.1</td>
<td>22.4</td>
<td>404.0</td>
<td>11.9</td>
</tr>
<tr>
<td>ì*</td>
<td>46.6</td>
<td>10.6</td>
<td>380.3</td>
<td>18.1</td>
<td>22.4</td>
<td>498.0</td>
<td>12.2</td>
</tr>
<tr>
<td>o</td>
<td>31.8</td>
<td>13.9</td>
<td>254.1</td>
<td>14.9</td>
<td>26.3</td>
<td>591.0</td>
<td>9.5</td>
</tr>
<tr>
<td>ø</td>
<td>39.8</td>
<td>13.9</td>
<td>420.5</td>
<td>15.3</td>
<td>26.3</td>
<td>283.0</td>
<td>11.1</td>
</tr>
<tr>
<td>u</td>
<td>25.8</td>
<td>11.3</td>
<td>182.9</td>
<td>14.2</td>
<td>26.7</td>
<td>166.0</td>
<td>9.9</td>
</tr>
<tr>
<td>f</td>
<td>34.9</td>
<td>5.6</td>
<td>146.1</td>
<td>10.9</td>
<td>17.1</td>
<td>196.0</td>
<td>9.2</td>
</tr>
<tr>
<td>θ</td>
<td>39.2</td>
<td>9.9</td>
<td>276.6</td>
<td>18.4</td>
<td>22.2</td>
<td>386.0</td>
<td>12.0</td>
</tr>
<tr>
<td>j</td>
<td>32.5</td>
<td>14.6</td>
<td>315.8</td>
<td>14.3</td>
<td>27.6</td>
<td>430.0</td>
<td>11.8</td>
</tr>
</tbody>
</table>

Configurations of [i] without teeth are denoted by "i*". The items corresponding to lines 3 and 4 pertain to the same replicas, as can be seen from their geometric parameters, i* being the configuration for which teeth have been removed from the replica. Lines 1 and 2 pertain to independent replicas.

Table 2 Average measured geometric parameters of the subject’s lips for the sustained sounds.

<table>
<thead>
<tr>
<th>Phoneme</th>
<th>A (mm)</th>
<th>B (mm)</th>
<th>S (mm²)</th>
<th>L (mm)</th>
<th>D (mm)</th>
<th>C (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>f</td>
<td>23.1</td>
<td>3.9</td>
<td>69.0</td>
<td>8.9</td>
<td>18.1</td>
<td>31.1</td>
</tr>
<tr>
<td>θ</td>
<td>39.0</td>
<td>8.9</td>
<td>255.2</td>
<td>17.5</td>
<td>23.0</td>
<td>21.6</td>
</tr>
<tr>
<td>s</td>
<td>38.6</td>
<td>6.6</td>
<td>201.1</td>
<td>17.6</td>
<td>20.2</td>
<td>23.2</td>
</tr>
<tr>
<td>j</td>
<td>33.2</td>
<td>9.1</td>
<td>232.9</td>
<td>13.1</td>
<td>21.9</td>
<td>29.9</td>
</tr>
<tr>
<td>ç</td>
<td>39.2</td>
<td>8.1</td>
<td>241.2</td>
<td>17.3</td>
<td>22.6</td>
<td>21.9</td>
</tr>
</tbody>
</table>

Fig. 5 Lip horn length L vs. lip width A for the subject.

Fig. 6 Lip horn length L vs. lip width A for the replicas (Configurations of [i] without teeth are denoted by "i*"; "i*" denotes the configuration that was tested both with and without teeth).

Finally, an interesting correlation is that of lip horn depth L with lip corner protrusion C (see Fig. 7). Abry and Boë® also reported such a correlation. In this case, the correlation between classes is high (−0.936), but it was verified that correlations within classes are still higher for the most protruded classes (−0.987 for [f], −0.977 for [f]), and lower
for the less protruded ones ($-0.699$ for $[c]$, $-0.416$ for $[s]$ and $-0.084$ for $[l]$). This shows the existence of a real physical phenomenon, which can be understood from the following imagery: the point of contact of the upper and lower lips moves forward as the lips come together, in the same way as the point of intersection of the two blades of a pair of scissors moves forward when these blades come together.

3.2.2 Acoustic parameters

Table 3 shows that acoustic and geometric lip areas are strongly correlated (correlation coefficient of 0.846), but that there is no correlation between acoustic and geometric lip horn depths (correlation coefficient of 0.11). The data are graphically represented in Figs. 8 and 9. For comparison, Motoki et al.'s data are also presented in the same figures.

![Fig. 7 Lip horn length $L$ vs. lip corner protrusion $C$ for the subject.](image_url)

![Fig. 8 Acoustic equivalent lip horn area $S_{\text{acoust}}$ vs. measured geometrical intralabial area $S_{\text{geom}}$ (Configurations of [i] without teeth are denoted by "i"). The data point for this study are marked with filled circles and those for the Motoki et al. (1988) by open circles. The regression lines computed for this study and for the Motoki et al. (1988) study are:

This study: $S_{\text{acoust}} = 1.096S_{\text{geom}} + 37.8$ (mm$^2$) ($R^2 = 0.716$)

Motoki et al.'s study: $S_{\text{acoust}} = 0.979S_{\text{geom}} + 24.4$ (mm$^2$) ($R^2 = 0.979$).

![Fig. 9 Acoustic equivalent lip horn depth $L_{\text{acoust}}$ vs. measured geometrical lip horn depth $L_{\text{geom}}$ (Configurations of [i] without teeth are denoted by "i"). The data point for this study are marked with filled circles and those for Motoki et al. (1988) by open circles.](image_url)

**Table 3** Correlation matrix for the geometric and acoustic parameters of the replicas (all configurations pooled).

<table>
<thead>
<tr>
<th></th>
<th>$A$</th>
<th>$B$</th>
<th>$S$ (geom.)</th>
<th>$L$ (geom.)</th>
<th>$D$</th>
<th>$A \cdot B$</th>
<th>$S$ (acoust.)</th>
<th>$L$ (acoust.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A$</td>
<td>1.000</td>
<td></td>
<td>.769</td>
<td>.737</td>
<td>-.393</td>
<td>.68</td>
<td>.445</td>
<td>.045</td>
</tr>
<tr>
<td>$B$</td>
<td>-.142</td>
<td>1.000</td>
<td>.503</td>
<td>.159</td>
<td>.944</td>
<td>.625</td>
<td>.668</td>
<td>.353</td>
</tr>
<tr>
<td>$S$ (geom.)</td>
<td>.769</td>
<td>.503</td>
<td>1.000</td>
<td>.697</td>
<td>.259</td>
<td>.977</td>
<td>.646</td>
<td>.214</td>
</tr>
<tr>
<td>$L$ (geom.)</td>
<td>.737</td>
<td>.159</td>
<td>.697</td>
<td>1.000</td>
<td>.013</td>
<td>.677</td>
<td>.374</td>
<td>.11</td>
</tr>
<tr>
<td>$D$</td>
<td>-.393</td>
<td>.944</td>
<td>.259</td>
<td>.013</td>
<td>1.000</td>
<td>.388</td>
<td>.495</td>
<td>.349</td>
</tr>
<tr>
<td>$A \cdot B$</td>
<td>.68</td>
<td>.625</td>
<td>.977</td>
<td>.677</td>
<td>.388</td>
<td>1.000</td>
<td>.858</td>
<td>.254</td>
</tr>
<tr>
<td>$S$ (acoust.)</td>
<td>.445</td>
<td>.668</td>
<td>.846</td>
<td>.374</td>
<td>.495</td>
<td>.858</td>
<td>1.000</td>
<td>.319</td>
</tr>
<tr>
<td>$L$ (acoust.)</td>
<td>.045</td>
<td>.353</td>
<td>.214</td>
<td>.11</td>
<td>.349</td>
<td>.254</td>
<td>.319</td>
<td>1.000</td>
</tr>
</tbody>
</table>

Major correlations are indicated in bold characters.
Table 4  Correlation matrix for the geometric parameters of the subject's lips for the sustained sounds (all configurations pooled).

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>S</th>
<th>L</th>
<th>D</th>
<th>A · B</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1.00</td>
<td>.715</td>
<td>.885</td>
<td>.94</td>
<td>.734</td>
<td>871</td>
<td>−.845</td>
</tr>
<tr>
<td>B</td>
<td>.715</td>
<td>1.00</td>
<td>.953</td>
<td>.572</td>
<td>.946</td>
<td>.96</td>
<td>−.389</td>
</tr>
<tr>
<td>S</td>
<td>.885</td>
<td>.953</td>
<td>1.00</td>
<td>.757</td>
<td>.925</td>
<td>.995</td>
<td>−.593</td>
</tr>
<tr>
<td>L</td>
<td>.940</td>
<td>.572</td>
<td>.757</td>
<td>1.00</td>
<td>615</td>
<td>.739</td>
<td>−.936</td>
</tr>
<tr>
<td>D</td>
<td>.734</td>
<td>.946</td>
<td>.925</td>
<td>.615</td>
<td>1.00</td>
<td>947</td>
<td>−.507</td>
</tr>
<tr>
<td>A · B</td>
<td>.871</td>
<td>.96</td>
<td>.985</td>
<td>.739</td>
<td>.947</td>
<td>1.00</td>
<td>−.596</td>
</tr>
<tr>
<td>C</td>
<td>−.845</td>
<td>−.389</td>
<td>−.593</td>
<td>−.936</td>
<td>−.507</td>
<td>−.596</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Major correlations are indicated in bold characters.

4. DISCUSSION AND CONCLUSION

The geometric measurements on the present corpus of sustained fricative consonants corroborate the "Laws for lips" established by Abry and Boë. Furthermore, although replicas obtained were too few to allow reliable statistical evaluation, it can be seen that geometric measurements of the replicas also tend to follow these laws. The quantitative discrepancies between the characteristics measured on the subject and on the replicas are likely due to the fact that the lip shape had somehow been distorted by the presence of the alginate material in the subject's mouth, and that the subject may not have sustained completely natural configurations. Nevertheless, results are qualitatively coherent. Moreover, the relationships between the acoustic and geometric characteristics of the replicas will not be altered by the fact that the dimensions and shapes of the replicas are not completely identical to those of the subject: it is thus possible to study the replicas as objects of themselves.

The results of the present study also confirm previous results obtained by Motoki et al. It appears that the measured acoustic equivalent area of the lips is linearly related to the geometric area, with a slope of 1.096, the correlation coefficient being 0.846. The dispersion in our data is higher than in those of Motoki et al., but note that the range of phonemes investigated in this present study is larger.

It was mentioned above that the teeth were not included in all the replicas. The case of [i] has allowed obtaining a first estimate of the influence of the teeth on the acoustic equivalent area. It appears from Table 1 that the acoustic area exceeds the geometric area by only 4.1 and 23.7 mm² for the two samples of [i] with teeth, whereas the acoustic area exceeds the geometric area by 69.4 and 117.7 mm² for the two samples of [i] without teeth. This confirms the contribution of teeth in decreasing the equivalent acoustic area at the lips. The decrease of the equivalent area has been also reported for the simplified geometric configuration of a circular tube with a constriction located near the exit. These results suggest that in the speech frequency range, the wave front would be deformed near the teeth. It can be thought of as a diffraction around a sharp edge. For simplified configurations, the equivalent length has been reported to increase when the constriction area decreases. However, comparing the equivalent acoustic lengths of [i] and [i*] pertaining to the same replica, it is difficult to confirm the former results.

Results concerning the equivalent acoustic length are shown in Fig. 9, and are coherent with those of Ref. 11. These results are however not satisfactory, in the sense that all attempts to fit the acoustic lip horn depth with any combination of the measured geometric parameters have failed. The general tendency is the reduction of the range of variation from the geometric domain (11–18 mm) to the acoustic domain (9–12 mm). This could be expected: indeed the flare of a horn will increase its resonance frequency, i.e. the effective length will be shorter. One may also intuitively conceive that a tube, with one end partly carved out, can appear acoustically shorter on the average: this is the case of the lip horn composed of the upper and lower lips, with an open interval between each other.

It was verified that the error on the equivalent acoustic length obtained by iterative computation is ± 0.5 mm, due to the average error on phase characteristics approximation. It should also be
recalled that the mechanical adjustment of the plexi-
glass tube to the different replicas was difficult, and
that it was especially difficult to ensure that the end
of the tube was really 1 mm behind the lip corner.
This uncertainty is large compared to the 3 mm
variation range measured, and shows that our es-
timation may have been affected by some important
errors. However, this range is so small that one
may, in a first approximation, consider the acoustic
equivalent length constant and equal to an average
of 11 mm. Further research is needed to refine this
characterization.

Finally a drawback of the acoustic measure-
ment method should be mentioned. The estima-
tion of the acoustic equivalence is based on the
measure of the radiation impedance: for front con-
figurations where a constriction is involved near the
labial horn, the radiation impedance is high, and
the reflection coefficient close to unity. In these
conditions, the accuracy of the determination is
rather poor. Therefore, other methods should be
investigated. Another limitation of this study is the
absence of mean flow. Indeed, it is known that the
presence of mean flow in a duct has an influence on
the end correction, as has recently been confirm-
ed.12)

In conclusion, “Laws for lips” by Abry and Boë3) have
been corroborated, as well as previous results
obtained by11) for acoustic equivalence. It appears
that the lip horn could be simply represented in a
first approximation by a single uniform tube with a
length of 11 mm and an area equal to the intralabial
area (possibly taking into account the coefficient
1.096). The lip corner position is thus the main
determinant of vocal tract length on the lip side.
These preliminary results need to be verified on a
larger corpus of configurations, and on more sub-
jects. The accuracy of the method should be im-
proved, both on the level of mechanical adjustments
precision and on the level of the acoustic radiation
impedance measurement procedure. Such an acous-
tic equivalence will be useful in the domain of acous-
tic and articulatory modeling of speech production.

ACKNOWLEDGEMENTS

This study was partially supported by the French
Centre National de la Recherche Scientifique, by the
Japanese Ministry of Education Grant No. 05044072
(International coresearch), and by TECNOVA
Corporation (Sapporo). We are grateful to many
people who helped us: Pierre Chardon, Kiyoharu
Kamada, Daniel Pelloux, and all the students at the
Research Institute for Electronic Science. Our
thanks also go sincerely to Christian Abry, Christian
Benoi, and Avraham Hirshberg, for their very
helpful comments on a draft version of this paper.
We are also very much indebted to Rudolph Sock
who corrected the English version of this paper.

REFERENCES

1) K. N. Stevens and A. S. House, “Development of
a quantitative description of vowel articulation,”
2) G. Fant, Acoustic Theory of Speech Production
3) G. Fant, “Vocal tract area functions of Swedish
vowels and a new three-parameter model,” Proc.
1992 Int. Conf. Spoken Language Processing, Banff,
Canada, Vol 1, paper Fr. 1AM. 3.1, 807-810 (1992).
4) L. J. Boë, P. Perrier, and G. Baill, “The geometric
vocal tract variables controlled for vowel produc-
tion: proposals for constraining acoustic-to-
5) C. Abry and L. J. Boë, “Laws for lips,”
6) C. Benoît, M. T. Lalouche, T. Mohamadi, and C.
Abry, “A set of French visemes for visual speech
synthesis,” in Talking Machines: Theories, Models
and Designs. G. Baill, C. Benoît et T. R. Sawallis,
Eds. (Elsevier Science Publications, Amsterdam,
7) J. L. Flanagan, Speech Analysis, Synthesis and Per-
ception (Springer Verlag, Berlin/Heidelberg/New
8) P. M. Morse, Vibration and Sound (McGraw-Hill,
New York/Kogakusha, Tokyo, 1948).
9) K. Motoki, N. Miki, and N. Nagai, “On the in-
fluence of the shape of radiation part upon the radia-
for the radiation part in consideration of real utter-
ance,” Trans. IECE Jpn. J60-A, 1269-1276
(1986).
11) K. Motoki, N. Miki, and N. Nagai, “Measurement of
12) P. Badin, “Fricative consonants: acoustic and X-ray
13) C. H. Shadle, P. Badin, and A. Mouliner, “To-
wards the spectral characteristics of fricative con-
sonants,” Proc. XIth Int. Congr. Phonetic Sciences,
14) C. Scully, E. Castelli, M. Shirt, and E. Brearley,
“Analysis and simulation of a speaker's aerody-
amic and acoustic patterns for fricatives,” J.


