Control of the vocal tract when experienced saxophonists perform normal notes and overtones

Tokihiko Kaburagi1,*, Ayame Kato2, Yuri Fukuda3 and Fumiaki Taguchi4

1Faculty of Design, Kyushu University, 4–9–1 Shiobaru, Minami-ku, Fukuoka, 815–8540 Japan
2Institute of Technology, Shimizu Corporation, 3–4–17 Etchujima, Koto-ku, Tokyo, 135–8530 Japan
3Fujitsu Kyushu Network Technologies Ltd., 2–2–1 Momochihama, Sawara-ku, Fukuoka, 814–8588 Japan
4Graduate School of Design, Kyushu University, 4–9–1 Shiobaru, Minami-ku, Fukuoka, 815–8540 Japan

(Received 10 June 2020, Accepted for publication 23 October 2020)

Abstract: Acoustic measurement of the vocal tract proved that experienced saxophonists tune their vocal tract during advanced performances to effectively influence the vibration frequency of the reed (Scavone et al., J. Acoust. Soc. Am., 123, 2391–2400 (2008); Chen et al., J. Acoust. Soc. Am., 129, 415–426 (2011)). To understand how the shape of the vocal tract is altered, the vocal tracts of experienced saxophonists were scanned in three dimensions with magnetic resonance imaging while they played the instrument using different pitches with normal and overtone techniques. The scanned images demonstrated that the tongue was located posteriorly in the vocal tract for low notes; however, it moved forward when the participants produced overtones. The input impedance was then calculated for the players’ air columns, including both the supra- and sub-glottal tracts, using an acoustic tube model. When the tongue moved forward to produce overtones, both the frequency and amplitude of the second impedance peak increased, suggesting an effective acoustic influence of the vocal behavior on the vibrating reed. The first impedance peak was less variable, regardless of the significant change in the vocal-tract shape for different notes.

Keywords: Saxophone, Overtone, Vocal tract, Input impedance, MRI

1. INTRODUCTION

For woodwind instruments, the vibration of the reed dynamically alters the cross section of the air channel, through which the stream of air from the lungs passes, resulting in a periodic change in volume flow at the narrow orifice between the reed and mouthpiece. Because of the temporal change in the airflow at the outlet of the orifice, acoustic pressure is generated, which is the source of the sound that drives the instrument air column and outputs the instrument sounds. When the reed is mechanically driven by the static pressure in the player’s mouth, the acoustic field formed inside the mouthpiece near the reed also affects the reed motion. This acoustic field reflects the acoustic characteristic of the instrument air column. The self-sustained oscillation of the reed arises near the resonance frequency of the instrument, because of the interaction between the elastic body of the reed, static lung pressure, and acoustic mouthpiece pressure. The basic sound generation mechanism of woodwind instruments was understood by investigating the autonomous vibration mechanism of the reed as a class of pressure-controlled valves [1,2] and the acoustic input impedance of the instrument air column [3].

However, during instrument performances, an acoustic field can also be created in the player’s vocal tract. The vocal tract constitutes an acoustic resonator, connected to the instrument via the reed generator. At frequencies near resonances of the instrument air column, the input impedance of the vocal tract may be smaller in magnitude than that of the instrument. Therefore, in many situations, the acoustic effect of the vocal tract can, to a first approximation, be ignored; this prediction was supported experimentally by Backus [4]. However, Benade [5] demonstrated that the vocal tract can be approximately...
connected to the instrument air column in series. Fritz and Wolfe [6] measured the input impedance of clarinetists’ vocal tracts under normal playing conditions, when the players mimed the performance and kept the vocal tract unchanged during measurements. Additionally, Chen et al. [7] measured the vocal-tract impedance acoustically, when the players actually blew into the instrument. For this method, pressure signals produced by the vibrating reed were eliminated post measurement by smoothing the measured impedance data. These studies showed that the vocal tract can be adjusted in specific performances, such as for pitch bending, and the magnitude of the vocal-tract impedance was comparable to that of the instrument. Based on these findings, they concluded that the resonance characteristics of the vocal tract could influence, select, or control the frequency of reed vibrations.

For the saxophone, Scavone et al. [8] measured the impedance ratio between the upstream and downstream of the reed generator at fundamental and harmonic frequencies of the vibrating reed. They demonstrated that the vocal-tract impedance could be greater than the bore impedance at the blowing frequency, during pitch bending; however, the complete impedance characteristics of the vocal tract were not determined in this study. In contrast, Chen et al. [9] conducted a more direct impedance measurement of the vocal tract during an advanced performance of the tenor saxophone, such as playing in the altissimo register, bugling, and pitch bending, with the method used for clarinetists [6,7]. The amplitude of the impedance peak of the tenor saxophone decreases more rapidly with frequency than that of the clarinet. Moreover, they showed that a strong resonance, created in the vocal tract of an experienced player, could compensate for the bore impedance, so that the vibrating frequency of the reed generator can be determined. Li et al. [10] used the same measurement method to examine how the vocal-tract impedance can influence the sound spectrum of the saxophone.

Although the acoustic significance of the vocal tract has been confirmed for the clarinetists and saxophonists through previous studies, the actual physiological control of the vocal tract has not been extensively studied. During performance, the lips should hold the mouthpiece and reed; therefore, the lip posture is less variable, and thus, less able to tune the vocal tract impedance. The position of the lower jaw is also less variable. Thus, the tongue is the most plausible articulator to change the configuration and impedance spectrum of the vocal tract. Among different observation methods of the vocal tract, magnetic resonance imaging (MRI) is a safe, noninvasive imaging technique for humans that is capable of obtaining a considerable amount of scan data [11,12]. A previous study that conducted volumetric scanning of the vocal tract reported that the cross-sectional area function could be accurately estimated [13,14]. The area function data have been reported to be useful for examining the acoustic characteristics of the vocal tract, using an acoustic tube model, including energy losses on the surface of the tract [15]. Such morphological and acoustic analysis of the vocal tract has been performed for a trumpet player [16], revealing that the method is effective for investigating the controlling strategy of the player’s vocal tract.

To observe the actual geometry of the vocal tract during performances by experienced saxophonists, we conducted an experiment to observe the shape of the vocal tract in three dimensions using MRI. Because the MRI device produces a strong magnetic field, the actual saxophone instrument could not be used; therefore, we connected a flexible vinyl tube into the mouthpiece and used it in place of the instrument in the MRI measurement. The impedance characteristic of the instrument was then fixed, and the vocal tract was scanned while participating players produced three notes with the instrument, i.e., the standard note and two overtones. The production of overtones has been examined for the bugling technique used for saxophones [8,9], indicating the involvement of the vocal tract in selecting the resonance of the instrument, especially for high notes, for which the magnitude of the instrument resonance decreases. Using the MRI measurement, we clarify how the shape of the vocal tract is altered to select the note, and examine the relationship between the configuration and impedance spectrum of the vocal tract. The opening of the glottis causes the loss of acoustic energy; hence, it is an important factor when determining the amplitude and bandwidth of vocal-tract resonances. We therefore examine how the glottal opening can influence the impedance spectrum of the vocal tract.

This paper is organized as follows: In Sect. 2, the instrument used in this study is explained, and the spectrum of normal and overtone sounds is analyzed. Section 3 provides a description of the image acquisition technique and the results of the morphological analysis. In Sect. 4, the method to calculate the input impedance of the vocal tract is described, and the numerical results are presented. Finally, Sect. 5 provides a discussion on our findings and the conclusions that can be drawn from the presented results.

2. SPECTRAL ANALYSIS OF INSTRUMENT SOUND

Prior to observing the vocal tract, we recorded the instrument sounds of the normal notes and overtones performed by a professional saxophonist. Because the MRI device produces a strong magnetic field, metallic materials (and hence the actual saxophone) could not be used. We therefore connected a flexible vinyl tube into the mouth-
piece of the saxophone for our experiments. The lengths of the uniform tube were 18 cm (short condition) and 28 cm (long condition), without the length of the tube tip inserted into the mouthpiece, and the internal radius of the tube was 0.5 cm. These two tubes of different lengths were used to increase the number of blowing frequencies to which the resonance of the vocal tract was adjusted, and the relationship between the blowing and resonance frequencies for a variety of blowing conditions was examined. The length of the short tube was approximately the same as that of the neck of the actual instrument. The ligature was made of synthetic fibers. The saxophonist blew into each instrument, producing three different musical notes, i.e., the lowest tone obtained by a normal playing and mid and high tones obtained by the overtone technique. The acoustic signal was measured using a microphone (Brüel & Kjær, type 4191) positioned at approximately 100 cm from the open end of the instrument. Each signal was amplified, using a microphone amplifier (Brüel & Kjær, type 2669 and NEXUS conditioning amplifier), and then, digitally saved to a computer at a sampling rate of 16 kHz with 16-bit resolution. The measurements were conducted in a soundproof room.

Figure 1 shows the spectrum (relative magnitude) of the recorded sound signals, where the left and right columns, respectively, correspond to the short and long tubes. Spectral analysis was performed using a recorded signal for a duration of 500 ms for each blowing condition, where the hamming window was used to extract a stable signal segment. The maximum frequency was set to 6 kHz in the figure to obtain a clear presentation of spectral patterns for every blowing condition. From top to bottom, the fundamental frequency for the short tube was 266, 884, and 1,362 Hz, approximately corresponding to the notes of C4, A5, and F6, respectively. For the long tube, the fundamental frequency was 200, 650, and 1,122 Hz, approximately corresponding to G3, E5, and C#6, respectively. Except for the high note of the short tube, the fundamental component was the most prominent. The figure shows that the player was able to produce three different tones for each instrument. Furthermore, it indicates that the three resonance modes for each instrument were switched by intentionally changing the blowing state, such as the mechanical pressure of the lower lip acting on the reed as well as the shape and acoustic property of the vocal tract, which can influence the reed motion.

The fundamental frequency of both tubes was within a moderately low frequency range of the alto saxophone when the normal playing technique was used. The frequencies of performed overtones were not extremely high if the input impedance of the player’s vocal tract should be tuned to perform these frequencies because our previous study showed that the frequency of the 2nd impedance peak can be as high as approximately 1,800 Hz [17]. Based on these preliminary results, we considered that the instruments were adequate to examine the players’ state during the performance of overtones. In addition, we were able to examine vocal tract control for various pitches by using two tubes of different lengths.

3. MORPHOLOGICAL ANALYSIS

3.1. Imaging of the Vocal Tract

Observation of the vocal tract shape is important to understand the mechanism underlying the distinction between normal and overtone playing techniques. Thus, in this study, we used MRI to non-invasively image the vocal tract. The participants were two professional saxophonists and they are denoted by A and B hereafter. Player A also participated in the audio recording experiment presented in the previous section. Sectional images of the vocal tract were first scanned in three dimensions while each participant produced three notes (low, mid, and high), using the short and long uniform tubes attached to the same mouthpiece. A guide tone was presented to the participant, which corresponded to the note of C4, A5, and F6 for the

Fig. 1 Magnitude spectra of the recorded sound signals performed by a saxophonist using the short (left) or long (right) uniform tube attached to the saxophone mouthpiece.
short tube, respectively, and G3, E5, or C#6 for the long
tube, respectively. The scanning started just after the
participant began to produce the target instrument sound.
The shape of the upper and lower teeth was then measured
independently because they are as transparent as air in
standard vocal-tract scanning. The oral shape was scanned
using the usual scanning procedures, while the participants
bit their tongues softly so that the outline of the teeth could
be distinguished in the images. In addition, the oral shape
was scanned when the tongue tip was attached to the
mouthpiece to determine the position of the edge of the
mouthpiece, i.e., the effective inlet of the vocal tract when
each player blew into the instrument.

Each slice of the volumetric data was measured in the
sagittal plane. The slice thickness was 3 mm for the vocal
tract during playing and 1 mm when observing the shape of
the teeth and mouthpiece edge. The measurements pro-
duced 24 slices for the vocal tract and 88 slices for the teeth
and mouthpiece edge. Obtaining a set of volumetric data
required a scanning time of approximately 30–40 s, and the
participant was instructed to maintain as much stability in
the configuration of the vocal tract as possible during the
scanning. The resolution of each slice was $512 \times 512$
pixels, corresponding to a scanning area of $256 \times 256$
mm$^2$. The measurement was made by the ATR-Promotions
Inc. (Kyoto, Japan).

Figures 2 and 3 show the midsagittal images of the
vocal tract during the production of each musical note. The
main speech articulators, including the tongue, lips, and
soft palate, can be clearly observed in the images. The
surfaces of both lips were flattened, and in particular, the
lower lip bended into the mouth to support the reed. The
posture of the lips and position of the lower jaw were less
variable; therefore, the tongue was the most important
articulator used to change the configuration and acoustic
characteristic of the vocal tract. For participant A, the
blowing frequencies during the MRI measurement were
285, 894, and 1,473 Hz for the low, mid, and high notes,
respectively, using the short tube and 201, 658, and
1,105 Hz, respectively, using the long tube. For participant
B, the blowing frequencies were 286, 867, and 1,303 Hz,
respectively, using the short tube and 197, 630, and
1,077 Hz, respectively, using the long tube. The blowing
frequency of each playing condition was not the same
between these two participants, and it was different from
that measured during the audio recording. This discrepancy
could have been caused by strong noises generated by the
MRI device during the scanning. The noise was composed of
multiple frequency components, which made it difficult
for the participant to hear their own blowing sound,
resulting in the lack of accurate auditory feedback.

For participant A (Fig. 2), the tongue was located
slightly towards the back of the tract for the low note, using
both the short and long tubes. Then, the vocal-tract
constriction was made between the tongue and soft palate.
For the mid and high notes, the tongue was positioned
more anteriorly, and the vocal tract was constricted
between the tongue and hard palate. When the degree of
this constriction is compared between the blowing fre-
quencies of 658, 894, and 1,105 Hz, the figure indicates
that the tongue approached the hard palate as the blowing
frequency increased. For the highest note, the tongue
was positioned most anteriorly, to form the constriction
between the tongue and gums.

For participant B (Fig. 3), the overall midsagittal shape
of the head was significantly different from participant A.
Here, it is important to note that the image resolution was
the same for both participants. However, the figure shows a
similar general tendency, i.e., that the tongue moved
forward as the blowing frequency increased. The deforma-
tion of the tongue was noticeable for the change from the
low to mid note. As a result, a clear vocal-tract constriction
was made between the tongue and hard palate for the mid
and high notes.
3.2. Cross-sectional Area Function

To investigate the acoustic characteristic of the players’ vocal tracts, i.e., the input impedance observed from the mouthpiece edge, we estimated the cross-sectional area of the vocal tract from each set of volumetric images. The images for the upper and lower teeth were first superimposed on the vocal-tract images, based on an affine congruent transformation [14]. Before applying the transformation, the position of the teeth was manually marked on the images. When the transformation was performed and the marked teeth area was mapped onto the tract images, the correlation coefficient between the teeth and tract images was calculated for the marked area. The optimal affine transformation was then determined, corresponding to the largest correlation coefficient.

The area function was then obtained using the vocal-tract images with the teeth superimposed. First, the inferior and superior outlines of the vocal tract were traced by sight on the midsagittal plane. The center line of the tract, representing the axis of wave propagation, was determined such that the distances from each node point on the center line to the two outlines were equal. Here, the anterior end of the center line was positioned at the edge of the mouthpiece; this point was determined from MRI images obtained when the participant attached the tip of the tongue to the mouthpiece. The other end of the center line was positioned at the center point of the glottis. The vocal tract was then divided into 40 sections of equal length along the center line. Finally, the cross-sectional area in the direction perpendicular to the center line was obtained for each section. The length of the vocal tract was also determined by the length of the center line.

The obtained cross-sectional area function data are plotted in Figs. 4 and 5, where the ordinate is the cross-sectional area and the abscissa is the vocal-tract section from the glottis to the mouthpiece edge.

![Fig. 3 Midsagittal image of the vocal tract when participant B blew into the instrument using the short (left) or long (right) uniform tube. The frequency shows the blowing pitch during the MRI measurement.](image)

![Fig. 4 Vocal-tract area function of participant A obtained from the volumetric MRI for the short (left) and long (right) tubes. The ordinate is the cross-sectional area, and the abscissa is the vocal-tract section from the glottis to the mouthpiece edge.](image)

The estimated lengths of the vocal tracts for participants A and B, respectively, were 14.9–16.0 cm and 16.0–18.2 cm as shown in Table 1.
4.1. Input Impedance of the Vocal Tract with the Glottal and Sub-glottal Sections

The acoustic characteristics of the vocal tract were analyzed using cross-sectional area function data. A frequency-domain acoustic tube model [15], which effectively accounts for the energy losses along the wall of the vocal-tract tube, was used for the analysis. Because the glottis is constantly open during playing, so that the flow of air from the lungs passes smoothly through the glottal orifice, we should consider the connection between the vocal tract, glottis, and sub-glottal airway, including the trachea, bronchi, and lungs, which forms an acoustic resonator.

To examine the influence of the glottal opening on the vocal-tract input impedance, we estimated the area and thickness of the glottis was uniformly shortened by a factor of 0.941 [19].

For the sub-glottal airway, the morphological data were obtained from a study by Weibel [18]. The airway can be classified into conductive, transitory, and respiratory zones, where the cross-sectional area is relatively constant along a section approximately 20 cm below the glottis, but increases rapidly in the remainder of the conductive zone and in the transitory and respiratory zones. The branching of airways was not directly considered here, and the cross-sectional area for each stage was summed to form a single uniform tube [19]. The resulting area function data were then used to calculate the input impedance of the vocal tract, where the total length of the sub-glottal airway was uniformly shortened by a factor of 0.941 [19].

For the vocal tract sections, from the mouthpiece edge to the glottis, defined as 1, 2, ..., N, and the sub-glottal sections, from the glottis to the lungs, defined as 1, 2, ..., M, the propagation matrix for the vocal tract is given by

\[
K_v = \left( \begin{array}{c} A_v \\ C_v \\ D_v \end{array} \right) = \prod_{i=1}^{N} \left( \begin{array}{cc} A_{vi} & B_{vi} \\ C_{vi} & D_{vi} \end{array} \right).
\]

the propagation matrix for the sub-glottal airway by

\[
K_s = \left( \begin{array}{c} A_s \\ C_s \\ D_s \end{array} \right) = \prod_{i=1}^{M} \left( \begin{array}{cc} A_{si} & B_{si} \\ C_{si} & D_{si} \end{array} \right)
\]

and the propagation matrix for the total airway by

\[
K = \left( \begin{array}{c} A \\ C \\ D \end{array} \right) = \left( \begin{array}{c} A_v \\ C_v \\ D_v \end{array} \right) \left( \begin{array}{c} A_s \\ C_s \\ D_s \end{array} \right)
\]

Matrix components for the vocal tract \((A_{vi}, B_{vi}, C_{vi}, \) and \(D_{vi})\) are determined by the length \(L_i\) and cross-sectional area \(S_i\) for the \(i\)th section using the acoustic tube
The matrix components for the sub-glottal airway, $A_{si}$, $B_{si}$, $C_{si}$, and $D_{si}$, including the glottal tube, were also calculated using the same acoustic tube mode [15]; however, the wall loss is greater for the sub-glottal airway than for the vocal tract, owing to an effective increase in the wall area. This effect was approximated by changing the parameter value of the tube model to $a = 130\pi$ rad/s, $b = (30\pi)^2$ (rad/s)$^2$, $c_1 = 4$ rad/s, $\omega_0^2 = (406\pi)^2$ (rad/s)$^2$, and $j = \sqrt{-1}$. $\rho$ is the air density and $c$ is the speed of sound. Here, $\rho$ was set to 1.184 × 10$^{-3}$ g/cm$^3$ and $c$ to 34,630 cm/s.

The input impedance of the total airway, including the vocal tract and sub-glottal airway, $Z$, was then calculated as

$$Z = \frac{AZ_p + B}{CZ_p + D},$$

where $Z_p = \rho c^2/(j\omega V)$ is the terminal impedance, modeled as an air volume ($V$) of 500 ml [20]. In addition, the input impedance ($Z_V$) of the vocal tract can be calculated as follows, under the assumption of complete glottal closure:

$$Z_V = -\frac{A_V}{C_V}.$$

Note that the sign for the traveling wave of the volume velocity, from the lungs to the mouthpiece, is positive.

### 4.2. Numerical Results

Figures 8 and 9 show the magnitude of the calculated input impedance. The area function data (Figs. 4 and 5) and vocal-tract lengths (Table 1) were used for the calculation of the vocal tract part. The glottis was modeled by a single uniform tube, with an area and length of 0.22 cm$^2$ and 0.25 cm, respectively, for participant A and 0.24 cm$^2$ and 0.30 cm, respectively, for participant B. The sub-glottal airway was also incorporated, as explained.
Previously, the input impedance was presented in these figures under two conditions, i.e., with (red line) and without (blue line) the sub-glottal airway, following the mathematical expression $Z$ and $Z_v$, given in Eqs. (4) and (5). For the condition without the sub-glottal airway, the glottis was assumed to be closed and the impedance was determined by the vocal tract section.

The figures indicate that the numerical results, particularly for participant B (Fig. 9), agreed well with the actual blowing frequency (broken vertical line) for both the short (left) and long (right) tubes. For the low note, for both tubes, the blowing frequency was in the vicinity of the first, i.e., the lowest, impedance peak. As the blowing frequency increased to the mid and high notes, the frequency for the second impedance peak also increased and the blowing frequency was very close to this peak. When comparing the result for the mid or high note between the tubes, the frequency of the 2nd impedance peak was higher when the participant used the short tube. This was in accordance with the higher blowing frequency of the short tube. Although the postures of the tongue shown in Fig. 3 are apparently similar for both tubes for each of the mid- and high-tone conditions, the impedance spectrum in Fig. 9 was apparently different and the frequency of the 2nd impedance peak was finely controlled based on the blowing frequency.

It is interesting to note that the frequency and magnitude of the first impedance peak were approximately the same among the three notes, although the participant significantly changed the vocal tract for each note (see Figs. 3 and 5). Another specific feature is the relative magnitude between the first and second peaks. For the low note, the magnitude of the second peak was only slightly greater than, or even smaller than, that of the first peak; however, it was much greater for the mid and high notes. Such prominence of the second peak may be important for producing mid and high notes when playing overtones. In addition, note that for the mid and high tones, the blowing frequency was consistently lower than the frequency of the second peak.

Moreover, a similar change in the relative magnitude between the first and second impedance peaks can be observed for participant A (Fig. 8), where the second peak was the most prominent for the mid and high notes, except for the high note when the short tube was used. However, the relationship between the blowing frequency and frequency of the prominent impedance peak was less consistent for this participant, especially for the mid note using the long tube, for which the second peak frequency was significantly higher than the blowing frequency.

For both participants, the frequency of the first impedance peak varied less irrespective of the dynamic deformation of the vocal tract, depending on the note. This suggests that the control of the frequency and magnitude of the second peak is the key factor to achieve the playing of overtones. Additionally, the control of the vocal tract generally corresponded with the change in the blowing frequency; thus, the acoustic field upstream of the reed could provide effective feedback to the oscillating reed.

Furthermore, the effect of the coupling between the vocal tract and sub-glottal airway can be observed from the figures. When the coupling was considered, the acoustic characteristic of the sub-glottal airway affected the impedance value, as follows. First, the frequency for the first peak of $Z$ (impedance with coupling) was greater than the corresponding peak frequency of $Z_v$ (impedance without coupling). The magnitude of the impedance peak tended to decrease and additional peaks and dips appeared with the effect of coupling. As a result, the effect of the coupling on the second impedance peak was complicated. For the mid and high notes, the frequency for the second peak of $Z$ did not substantially change from that of $Z_v$, typically for participant B. For the low note, the coupling weakened the second peak appearing in $Z_v$ and as a result, the first peak was more emphasized than the second one.

Next, Figs. 10 and 11 show the phase angle of the input impedance. For the mid and high tones played by participant B, the blowing frequency was consistently below the frequency of the second impedance peak; the phase angle was positive (at approximately $70^\circ$–$80^\circ$) at the blowing frequency. This phase property is known to be effective for the maintenance of the self-sustained oscillation of the reed [2]. For the low note of the short and long tubes played by both participants, the top plots of both figures indicate that the phase angle was also positive at the
blowing frequency when the vocal tract and the subglottal airway were coupled because the frequency of the first peak increased due to coupling. For both participants, it was more apparent observed when a long tube was used. The mid and bottom plots in Fig. 10 show that the results for the mid and high notes played by participant A were complicated to interpret; however, when the effect of coupling was considered, the phase angle was at least positive, except for the high tone using the long tube.

5. SUMMARY AND CONCLUSION

A morphological and acoustic analysis was performed to examine the possible role of a player’s vocal tract on the production of saxophone overtones. We constructed two instruments by connecting a uniform tube, of 18 or 28 cm in length, to a standard mouthpiece, for use in the MRI device. The vocal tracts of two professional saxophonists were scanned in three dimensions, using MRI during the production of three musical notes: the lowest one was obtained via the normal playing technique and the mid and high notes via the overtone playing technique. The image data revealed that the tongue constricted the vocal tract in different ways, depending on the note produced. Moreover, the cross-sectional area function was determined from the vocal-tract images to calculate the input impedance, observed from the mouthpiece edge. The main results obtained in this study can be summarized as follows.

First, the scanned images showed that the tongue was positioned towards the back of the tract for the lowest note and more anteriorly for the mid and high notes. Vocal-tract constriction was observed between the tongue and soft palate for the low note, and a narrow channel was formed between the tongue and hard palate or between the tongue and gums, for the mid and high notes. These results were also confirmed by the cross-sectional area function data. The results were in agreement with the study conducted on a trumpet player [16]; however, a more significant deformation of the tongue was observed for the saxophonists in this study.

Second, the calculated input impedance of a participant corresponded with the blowing frequency; the frequency of the second impedance peak was consistently slightly below the blowing frequency of the overtones, for which the phase angle of the input impedance was approximately 70°–80°. This result was in good agreement with that measured for saxophone bugling [9]. In addition, the magnitude of this peak was greater for the overtone technique than for normal playing. Another participant only partly showed the same strategic control of the vocal tract. This participant possibly used other blowing parameters, for example, the effective mechanical property of the reed, to perform the overtone technique.

Third, while a change in the vocal-tract shape had a significant effect on the second impedance peak of the vocal tract, both the frequency and magnitude of the first impedance peak varied less for both the normal notes and overtones.

Based on these findings, the morphological control of the vocal tract and its acoustic consequence can be explained as follows. Morphologically, the tongue was positioned posteriorly for normal playing and anteriorly when playing overtones. When the tongue moved forward, the frequency of the second impedance peak and its magnitude simultaneously increased, so that this peak was more prominent than the first peak. As a result, the vocal tract could acoustically support the reed to perform overtones. However, the prominence of the second peak diminished when the tongue was located posteriorly, which may be suited for playing normal notes.

Methodologically, only MRI can noninvasively visualize the vocal tract in three dimensions. The actual instrument could not be used during measurements; however, the relationship between the morphological and acoustic characteristics of the vocal tract can be examined using high-resolution volumetric data, allowing the identification of the true nature of the acoustic mechanism underlying the performance of wind instruments. During the measurements, the participants could play the instrument with their desired volume in contrast to acoustic measurement methods [6,7,9,10], for which the participant should mime or play softly to obtain a high signal-to-noise ratio. While an ultrasound method can visualize the dynamic change in the tongue contour [21], MRI is also capable of imaging the vocal tract at a frame rate of several
dozen frames per second [22]. For our future studies, we intend to examine a larger number of participants and wider range of vocal-tract shapes for various musical notes, to confirm and generalize the present findings. Moreover, we plan to observe dynamic deformation patterns of the vocal tract during the performance of, for example, pitch bending, glissandi, and tonguing of woodwind instruments, with a real-time MRI.

Additional future work includes the computer simulation of instrument blowing by integrating models of the reed generator, instrument bore, and vocal tract [23]. We have already conducted a preliminary examination, where the model and parameter values of the reed were set, following the literature regarding the inverse estimation of reed parameters [24,25]. However, until now, we have not succeeded in reproducing the distinction between playing normal notes and overtones by simply switching the impedance spectrum of the vocal tract, determined by the current study, because the vibration frequency of the reed was under the strong influence of its natural frequency.

REFERENCES

[1] N. H. Fletcher, “Excitation in woodwind and brass instruments,” Acustica, 43, 63–72 (1979).
[2] N. H. Fletcher, “Autonomous vibration of simple pressure-controlled valves in gas flows,” J. Acoust. Soc. Am., 93, 2172–2180 (1993).
[3] J. Backus, “The input impedance curves for the reed woodwind instruments,” J. Acoust. Soc. Am., 56, 1266–1279 (1974).
[4] J. Backus, “The effect of the player’s vocal tract on woodwind instrument tone,” J. Acoust. Soc. Am., 78, 17–20 (1985).
[5] A. Benade, “Air column, reed and player’s windway interaction in musical instruments,” in Vocal Fold Physiology, I. Titze and R. Scherer, Eds. (The Denver Center for the Performing Arts, Denver, 1983).
[6] C. Fritz and J. Wolfe, “How do clarinet players adjust the resonances of their vocal tracts for different playing effects?,” J. Acoust. Soc. Am., 118, 3306–3315 (2005).
[7] J.-M. Chen, J. Smith and J. Wolfe, “Pitch bending and glissandi on the clarinet: Roles of the vocal tract and partial tone hole closure,” J. Acoust. Soc. Am., 126, 1511–1520 (2009).
[8] G. P. Scavone, A. Lefebvre and A. R. da Silva, “Measurement of vocal-tract influence during saxophone performance,” J. Acoust. Soc. Am., 123, 2391–2400 (2008).
[9] J.-M. Chen, J. Smith and J. Wolfe, “Saxophonists tune vocal tract resonances in advanced performance techniques,” J. Acoust. Soc. Am., 129, 415–426 (2011).
[10] W. Li, J.-M. Chen, J. Smith and J. Wolfe, “Effect of vocal tract resonances on the sound spectrum of the saxophone,” Acta Acust. United Ac., 101, 270–278 (2015).
[11] M. Rokkaku, K. Hashimoto, S. Imaizumi, S. Niimi and S. Kiritani, “Measurements of the three-dimensional shape of the vocal tract based on the magnetic resonance imaging technique,” Ann. Bull. RILP, 20, 47–54 (1986).
[12] T. Baer, J. C. Gore, L. C. Gracco and P. W. Nye, “Analysis of vocal tract shape and dimensions using magnetic resonance imaging: Vowels,” J. Acoust. Soc. Am., 90, 799–828 (1991).
[13] B. H. Story, I. R. Titze and E. A. Hoffman, “Vocal tract area function from magnetic resonance imaging,” J. Acoust. Soc. Am., 100, 537–554 (1996).
[14] H. Takemoto, T. Kitamura, H. Nishimoto and K. Honda, “A method of tooth superposition on MRI data for accurate measurement of vocal tract shape and dimensions,” Acoust. Sci. & Tech., 25, 468–474 (2004).
[15] M. M. Sondhi and J. Schroeter, “A hybrid time-frequency domain articulatory speech synthesizer,” IEEE Trans. Acoust. Speech Signal Process., 35, 955–967 (1987).
[16] T. Kaburagi, N. Yamada, S. Fukui and E. Minamiya, “A methodological and preliminary study on the acoustic effect of a trumpet player’s vocal tract,” J. Acoust. Soc. Am., 130, 536–545 (2011).
[17] T. Kaburagi, T. Shimizu and Y. Uezu, “A morphological and acoustic analysis of the vocal tract during the act of whistling,” Acoust. Sci. & Tech., 39, 198–206 (2018).
[18] E. R. Weibel, Morphometry of the Human Lung (Springer-Verlag, Berlin, 1963), pp. 136–140.
[19] K. Ishizaka, M. Matsuda and T. Kaneko, “Input acoustic-impedance measurement of the subglottal system,” J. Acoust. Soc. Am., 60, 190–197 (1976).
[20] van den Berg, “An electrical analogue of the trachea, lungs and tissues,” Acta Physiol. Pharmacol. Neerl, 9, 361–385 (1960).
[21] S. M. Lulich, S. Charles and B. Lulich, “The relation between tongue shape and pitch in clarinet playing using ultrasound measurements,” J. Acoust. Soc. Am., 141, 1759–1768 (2017).
[22] P. W. Ilitis, E. Schoonderwaldt, S. Zhang, J. Frahm and E. Altenmüller, “Real-time MRI comparisons of brass players: A methodological pilot study,” Hum. Mov. Sci., 42, 132–145 (2015).
[23] S. D. Sommerfeldt and W. J. Strong, “Simulation of a player-clarinet system,” J. Acoust. Soc. Am., 83, 1908–1918 (1988).
[24] V. Chatziioannou and M. van Walstijn, “Estimation of clarinet reed parameters by inverse modeling,” Acta Acust. United Ac., 98, 629–639 (2012).
[25] A. M. Arancón, B. Gazengel, J.-P. Dalmont and E. Conan, “Estimation of saxophone reed parameters during playing,” J. Acoust. Soc. Am., 139, 2754–2765 (2016).