

Vocal tract resonances and the sound of the Australian didjeridu (yidaki) I. Experiment^{a)}

Alex Z. Tarnopolsky

School of Physics, University of New South Wales, Sydney NSW 2052, Australia

Neville H. Fletcher

School of Physics, University of New South Wales, Sydney NSW 2052, Australia and Research School of Physical Sciences and Engineering, Australian National University, Canberra 0200, Australia

Lloyd C. L. Hollenberg

School of Physics, University of Melbourne, Melbourne, Vic 3010, Australia

Benjamin D. Lange, John Smith, and Joe Wolfe^{b)}

School of Physics, University of New South Wales, Sydney NSW 2052, Australia

(Received 8 August 2005; accepted 8 November 2005)

The didjeridu, or yidaki, is a simple tube about 1.5 m long, played with the lips, as in a tuba, but mostly producing just a tonal, rhythmic drone sound. The acoustic impedance spectra of performers' vocal tracts were measured while they played and compared with the radiated sound spectra. When the tongue is close to the hard palate, the vocal tract impedance has several maxima in the range 1–3 kHz. These maxima, if sufficiently large, produce minima in the spectral envelope of the sound because the corresponding frequency components of acoustic current in the flow entering the instrument are small. In the ranges between the impedance maxima, the lower impedance of the tract allows relatively large acoustic current components that correspond to strong formants in the radiated sound. Broad, weak formants can also be observed when groups of even or odd harmonics coincide with bore resonances. Schlieren photographs of the jet entering the instrument and high speed video images of the player's lips show that the lips are closed for about half of each cycle, thus generating high levels of upper harmonics of the lip frequency. Examples of the spectra of "circular breathing" and combined playing and vocalization are shown. © 2006 Acoustical Society of America. [DOI: 10.1121/1.2146089]

PACS number(s): 43.75.Fg, 43.75.Yy, 43.72.Ct [DD]

Pages: 1194–1204

I. INTRODUCTION

The word "didjeridu" (or "didgeridoo" in the popular literature) is an onomatopoeic Western name for a traditional instrument played in parts of Northern Australia and known to the Yolngu people of Arnhem Land as the yidaki. The Yolngu name "yidaki" will be used throughout this paper. It is unusual among wind instruments in that the pitch is only rarely varied: the interest in performance lies in spectacular, rhythmic variations in timbre, which are produced by the player's vocal tract. It is played using "circular breathing" to produce an uninterrupted sound: the player, traditionally a man, fills his cheeks and uses this reservoir to continue to play while simultaneously inhaling quickly through the nose (and bypassing the mouth at the soft palate) to refill the lungs with air. The differences between timbres produced by playing using the mouth cavity alone, while inhaling, and those produced using the complete tract, while exhaling, are usually unavoidable and are incorporated into the rhythmic variation in timbre that is idiomatic for the instrument and

that gives it its Western name. Different tongue positions have a strong effect on the sound spectrum. Sound files illustrating these effects are given at www.phys.unsw.edu.au/~jw/yidakididjeridu.html.

The yidaki is a member of the lip valve family. In this family, the playing frequency is usually close to that of one of the maxima in the impedance spectrum of the bore. The effect of variations in the player's vocal tract upon orchestral lip valve instruments is usual modest, because of their narrow bore and the shape of the mouthpiece, whereas in the yidaki it is the preeminent musical feature. The yidaki is therefore an ideal instrument in which to study the interaction among vocal tract, lips, and instrument.

Previous studies of the acoustics of the yidaki have considered the lip motion (Wiggins, 1988), the lip-bore interaction (Fletcher, 1983, 1996), numerical modeling of the lip motion (Hollenberg, 2000), the linear acoustics of the instrument (Amir and Alon, 2001; Amir, 2004), and the acoustics of the vocal tract of players miming playing (Fletcher *et al.*, 2001). A very brief report covering work related to that reported here has been given previously (Tarnopolsky *et al.*, 2005). That report used an instrument made in the traditional manner, and therefore of unknown geometry.

^{a)}A Brief Communication reporting related studies has been published in *Nature* (Tarnopolsky *et al.*, 2005).

^{b)}Author to whom correspondence should be addressed. Electronic mail: j.wolfe@unsw.edu.au

In what is now a standard model, Backus (1985) proposed that the acoustic impedance of the bore of a wind instrument Z_{inst} and that of the tract Z_{tract} act in series on the valve and on the air flow through it. An understanding of the interaction between the instrument and the vocal tract thus requires detailed knowledge of the acoustic impedance of the instrument, the acoustic impedance of the vocal tract, the vibratory behavior of the lip valve, the air jet entering the instrument, and their interaction.

It is relatively easy to determine the impedance of the instrument, particularly if only the maxima are important. However, the impedance of the vocal tract is much harder to measure, particularly during playing. The vibrating lips generate a sound signal that is transmitted both into the instrument and into the mouth. Consequently, the sound level inside the mouth of the player is very high. This makes it difficult to make measurements of the acoustic impedance of the vocal tract of someone playing the yidaki. In the past, we have made measurements of that impedance while players mimed playing (Fletcher *et al.*, 2001). However, producing a given mouth configuration in the absence of audible feedback is difficult, and it is not clear that yidaki players are capable of miming reliably, perhaps particularly with regard to the aperture of the glottis, of which most people are not conscious.

In this study, we report the development of a system that allows the measurement of the acoustic impedance of the vocal tract, just inside the player's lips, while he is playing the yidaki. We compare this with the spectrum of the sound produced. We also report the motion of the player's lips, using high-speed photography. These sets of observations are used to test a simple model that explains how the acoustic impedance of the tract affects the spectral envelope of the sound produced. We also analyze other features of idiomatic playing: "circular breathing" and vocalization.

A. The yidaki (didjeridu)

Traditionally, the material for a yidaki is selected by tapping suitably sized tree trunks to find one whose interior has been eaten by termites to provide a suitable central bore. It is cut to a desired length, the bore is cleaned and sometimes shaped further, and a ring of beeswax is fitted to the smaller end to make a comfortable seal for the player's lips. The outside is sometimes painted with traditional designs of cultural significance. The instrument is typically 1.2 to 1.5 m long (different cultural groups have different styles) and has an irregular bore, which is usually somewhat flared from about 30 to 50 mm at the blowing end to about 40 to 150 mm at the open end. Sealed at the lip end and open at the other, its lowest resonance is typically 50 to 80 Hz. Orchestral wind instruments usually have several bore resonances whose frequencies fall in harmonic ratios. Because of its shape, this is not usually the case for the yidaki: it is usually neither a cylinder nor a nearly complete cone and its resonances form a "stretched" quasi-harmonic series (Fletcher, 1996). Consequently, harmonics of the note being played only sometimes coincide with a bore resonance of the instrument. Higher or "overblown" notes near the fre-

quencies of the second or third bore resonances may be sounded briefly for contrast. This is not usual in traditional playing in Western Arnhem Land, though it is used in the East.

Because the bore is typically 30 to 50 mm in diameter at the smaller end, its characteristic impedance is lower by an order of magnitude than that of most of the members of the lip-driven musical instrument family such as the horn or trumpet. This and the rather rough walls of the bore imply that the magnitudes of the maxima in the yidaki's impedance spectrum are rather lower than those of other members of the wind instrument family. Some consequences have been discussed in previous papers (Fletcher, 1996; Amir and Alon, 2001; Fletcher *et al.*, 2001; Caussé *et al.*, 2004).

Because the shape of the bore is largely determined by termites and the shape of the tree trunk, the variation among these instruments is great. The purpose of the current study is to investigate the principles of operation, rather than the effects of different instrumental geometries (which is the subject of another study). For that reason, and to facilitate reproduction of the results reported here, two model instruments were used. For acoustical measurements, we used a cylindrical PVC pipe. For optical measurements, we used a pipe made of plexiglass with a square cross section. Experienced players reported that both model instruments played moderately well. Indeed instruments with a constant cross section, usually made of PVC pipe, are occasionally used in nontraditional musical contexts, particularly when a given pitch is required in order to play with other instruments.

B. Vocal tract-instrument interaction

There are a number of reports on the effects of the vocal tract on the sound on orchestral wind instruments (Elliot and Bowsher, 1982; Clinch *et al.*, 1982; Wolfe *et al.*, 2003), but the effects in such instruments are modest in comparison with those in the yidaki. These orchestral instruments have a narrow constriction in the mouthpiece and a smooth bore, which is typically only several mm in the mouthpiece. These features give the instruments an impedance spectrum with a series of maxima whose values exceed considerably those of the vocal tract. Consequently, there is only modest coupling between the two resonators (the vocal tract and the bore of the instrument). While the effect of the tract on timbre of orchestral instruments is large enough to interest composers (e.g., Berio, 1966; Erikson, 1969), it is small compared to the striking effects of the vocal tract on the timbre of the yidaki. In a previous paper (Fletcher *et al.*, 2001) we reported sound spectra, vocal tract configurations, and the impedance spectra of players miming the playing of the yidaki. However, these were not measured simultaneously during playing, so we were then unable to make quantitative comparisons among them.

C. Measurement of the vocal tract impedance during performance

The magnitude of the vocal tract effect in the yidaki makes it an ideal instrument upon which to study vocal tract

effects in general. For this purpose, we have adapted an impedance spectrometer described previously (Epps *et al.*, 1997) to allow us to make impedance measurements using an impedance probe placed just inside the player's lips, while he is playing. This situation requires several practical compromises. The sound due to the playing has comparable levels in the mouth and in the instrument. As this is "noise" for the purposes of measurement of the impedance spectrum, the signal-to-noise ratio is low. There is the further complication of a humid environment, which means that water-resistant or disposable microphones must be used. On the other hand, this study is concerned with relating the spectral envelope of the sound produced to the overall features of the impedance spectrum. Consequently, high-precision calibrated microphones are not required.

II. MATERIALS AND METHODS

A. Yidakis

For acoustic measurements, a "model yidaki" was made of cylindrical PVC pipe, length 1210 mm and inner diameter 30 mm. (It is referred to as "the pipe" or "the instrument" below.) For the optical measurements of the lip motion and for the flow visualization, a pipe with square cross section was made of plexiglass with glass panels for the optical pathway. It is 1220 mm long and the internal width is 38 mm.

B. Players

One of the players, BL, traditional name Wilamara, is a member of the Mara people of Roper River in Northern Australia, where he learned to play yidaki in the traditional style. LH is an Australian of European cultural background who has been playing the yidaki for 8 years. AT is an Australian of European cultural background who learned to play yidaki for the purposes of this study. LH's usual playing style has the instrument displaced laterally from the center of the lips. Neither he nor AT had trouble adapting to the presence of the impedance probe behind the lips. BL, who has played for the longest time and whose lip-instrument position is symmetrical, found the impedance probe disruptive, particularly for the high tongue position.

Players were asked to produce three different mouth configurations for recordings. One is called a high tongue drone (hereafter "high tongue"): the player holds the tongue close to the hard palate so that there is a constriction in the air passage between the throat and the lips. This produces a strong formant between about 1.5 and 2 kHz, whose frequency and amplitude may be varied by the performer. This sound is very common in yidaki performance. In another configuration, hereafter called "low tongue," the players were asked to play with the tongue low in the mouth and thus no lingual constriction. This configuration produces a sound without a strong formant and is used as a contrast to the high-tongue drone. In the third configuration, players inflated their cheeks and then expelled the air, while inhaling, as described above under "circular breathing." In a different series, they were asked to vocalize at harmonic intervals above the note they were playing on the yidaki.

To measure the static mouth pressure during playing, they were also asked to play with a range of loudness levels while a small tube connected the mouth cavity to a water manometer.

C. Measurements of impedance spectra of the instrument

An impedance spectrometer described previously (Smith *et al.*, 1997; Epps *et al.*, 1997) was adapted for this study. Briefly, a waveform is synthesized from harmonic components, amplified, and input via a loudspeaker and impedance matching horn to a narrow high-impedance tube leading to the item under test. This approximates an ideal source of acoustic current. It is calibrated by connection to a reference impedance, which is an acoustically quasi-infinite cylindrical pipe whose impedance is assumed to be real, frequency independent, and equal to its calculated characteristic impedance. From the coefficients of the spectrum of the measured sound in this calibration stage, a new signal is synthesized to produce a measured spectrum with frequency components of equal amplitude. This is used as the acoustic current source for subsequent measurements and the unknown impedance spectrum is calculated from the pressure components measured in measurement and calibration stages, taking into account the small, parallel admittance of the source. The bore diameter of the yidaki is larger than that of the instruments we have studied previously (Wolfe *et al.*, 2001) and consequently a lower impedance reference was required. The acoustically infinite cylindrical pipe used for calibration in this study had an internal diameter of 26.2 mm and a length of 194 m. Because the first curve in the pipe occurs at 40 m from the spectrometer and because any curves have a radius of 5 m or greater, the effects of reflections from these curves are expected to be negligible and this reference impedance should be purely resistive.

D. Measurements of impedance spectra in the vocal tract

The microphone recording the pressure inside the tract is exposed to high humidity and high steady pressure. For this reason, we used inexpensive electret microphones (Optimus 33-3013), which were replaced when necessary. It was necessary to attenuate the acoustic signal to avoid clipping or harmonic distortion in these microphones. To do this, we used the acoustic divider circuit shown in Fig. 1. At low frequencies, where attenuation is most important, the impedances of both pipes in the divider (including the radiation impedance associated with the open end of the tube) are essentially inertive, and the phase change along the length is small. At higher frequencies, there is a phase shift and a frequency-dependent gain. The microphone therefore records only a fraction of the pressure inside the mouth. This attenuation is inside the calibration loop for the impedance probe, so its frequency and phase response do not affect measurements. The source capillary has an inner diameter of 3.7 mm and the microphone tube an inner diameter of 1.5 mm. Both have a length of 35 mm. Measurements were made at a frequency spacing of 5.383 Hz from 0.2 to 3.0 kHz.

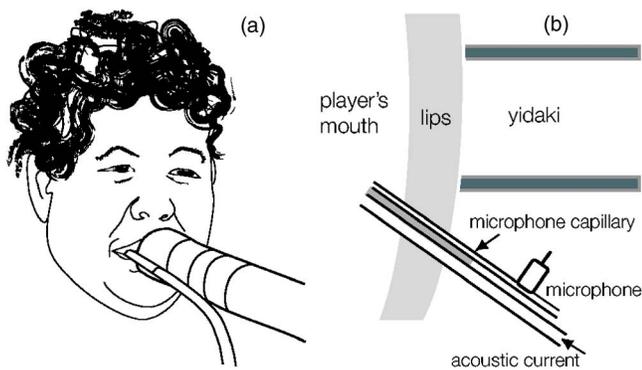


FIG. 1. The technique used to measure the impedance of the vocal tract during performance. The sketch shows the impedance probe inserted into a corner of the mouth. The schematic (top view—only approximately to scale) shows the geometry of the impedance probe and its location in the player's mouth. The microphone capillary is hidden in the figure at the left.

While the acoustic pressure acts upon a significant area of the lips and this determines their vibratory motion, the pressure of concern in the production of formants in the sound is that acting over the opening area of the lips. Measurements of acoustic impedance need to take this geometrical mismatch into consideration, since essentially there is an inertive correction, which may be either positive or negative, involved between this impedance and the plane-wave impedance normally measured in a pipe (Brass and Locke, 1997; Fletcher *et al.*, 2005). The probe used for measurement of the vocal tract impedance, shown in Fig. 1, has a narrow outlet for the acoustic current. The impedance spectrometer is calibrated on a quasi-infinite tube of diameter 26.2 mm, which is comparable with the size of the vocal tract. Consequently, errors due to this effect are small.

These measurements required the impedance probe shown in Fig. 1 (8 mm wide and 5 mm high) to be placed in the mouth during playing. Impedance measurements start after the player gives a signal that he is happy with the tongue position and the sound produced. The player then continues to play on one breath, typically for about 10 s. During this time an impedance measurement is made. A sample of the sound immediately following the impedance measurement is used to obtain samples of the radiated sound uncontaminated by that of the injected measurement signal.

E. Measurements of sound spectra

The output sound was recorded on digital audio tape at 44.1 kHz using an omnidirectional electret microphone, placed on the axis of the yidaki, at a distance of 125 mm from its end.

F. Measurements of the lip motion

The transparent yidaki with square cross section (mentioned above) had two mouthpiece configurations. For measurements of the lip motion, a round hole was cut in one side for the player's lips. To improve the image quality, sections of plexiglass on the side and end were replaced with two glass panels, 100 mm long and equal in width to the yidaki, inserted in the optical path at the mouth end, as shown in

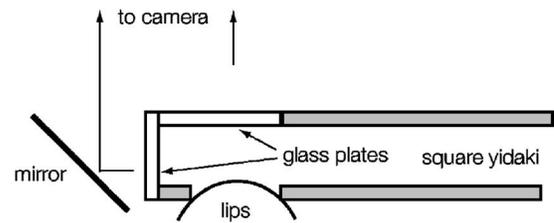


FIG. 2. The geometry of the square yidaki, seen from above, as configured for high speed photography.

Fig. 2. A mirror, mounted vertically on the end at 45°, allowed the single camera to record the plane and lateral image of the player's lips simultaneously. Before each recording, the glass panels were heated with warm dry air to prevent water condensation. A video camera running at 1000 frames per second was used to record the images.

G. Flow visualization of the jet motion

Images of the air jet inside the yidaki during playing were achieved using schlieren imaging, a nondestructive optical flow visualization technique, which is described elsewhere (Tarnopolsky and Fletcher, 2004).

This experiment also used the plexiglass yidaki. The schlieren technique depends upon refraction of light rays as they pass through regions with varying refractive index, usually provided by inhomogeneities in density. A curtain of higher density gas was produced by releasing carbon dioxide from a manifold on the outer side into the yidaki through a line of 11 holes of 1.5-mm diameter linking the manifold to the yidaki (see Fig. 3). During playing, the air jet passes through and is contaminated by the curtain of carbon dioxide and thus generates the necessary density gradient.

A light source that produced a single pulse of duration of about 0.2 ms was triggered electronically at a selected phase of the lips' opening (Tarnopolsky *et al.*, 2000). Stroboscopy of a steady, sustained playing gesture was used to obtain images covering one period of the lip's oscillation in time steps of 1 ms. The period of the lip oscillation was 14.3 ms for these experiments.

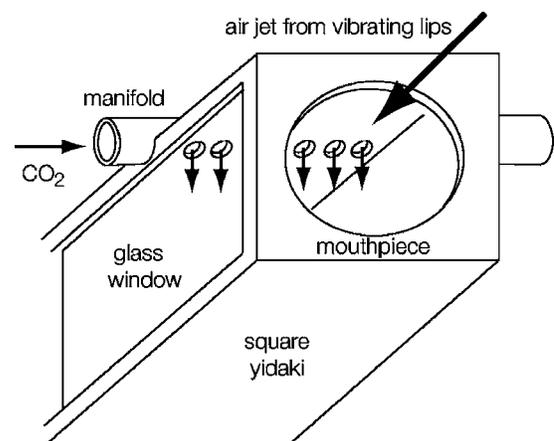


FIG. 3. The geometry of the square yidaki, as configured for flow visualization.

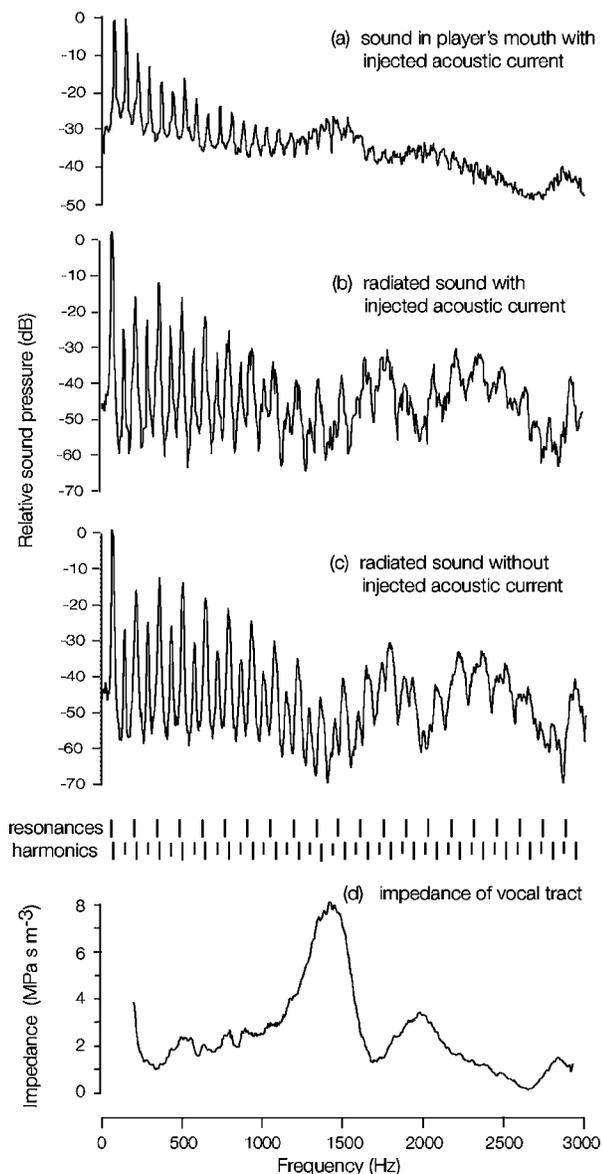


FIG. 4. A figure to illustrate the vocal tract impedance measurement while the subject is playing in the high tongue configuration. (a) shows the spectrum of the sound pressure level measured in the mouth due to both lip vibration and the injected acoustic current. (b) shows the radiated sound spectrum measured simultaneously with (a). (c) shows the radiated sound measured just after the impedance measurement i.e. without the injected acoustic current. (d) shows the impedance of the vocal tract that was derived from (a). The sound pressure levels shown in (a), (b), and (c) are normalized relative to their largest frequency component. The frequencies of the harmonics of the sound and of the resonances of the pipe are also shown with vertical dashes. The odd numbered harmonics are represented by longer dashes.

III. RESULTS AND DISCUSSION

A. Impedance of the player's vocal tract

Figure 4 shows how measurements of the acoustic impedance of the player's tract were made during performance and then processed. Figure 4(a) shows the spectrum of the signal recorded inside the mouth during a typical example of the high tongue configuration. The periodic vibration of the lips at about 70 Hz generates an acoustic signal that interacts with the impedance of the tract to produce a series of harmonics, which are seen at frequencies below about 1.5 kHz.

Inside the mouth, the higher resonances of the instrument (some of which lie close to odd harmonics of the lip motion) have only modest influence on the sound in the mouth. Hence, in Fig. 4(a), there is no systematic difference between odd and even harmonics. At frequencies above 1 kHz (the range of interest), the spectrum is increasingly dominated by the response of the vocal tract to the injected acoustic current. Here we see broad peaks or formants at about 1.5, 2.1, and 2.8 kHz. Because the injected current has been calibrated to have flow components with magnitude independent of frequency, these peaks correspond to maxima in the acoustic impedance spectrum of the vocal tract in this configuration.

Figure 4(b) shows the spectrum of the externally radiated sound produced by the yidaki, measured 125 mm from the end of the instrument. This shows the strong harmonics of the instrument's sound. Odd harmonics dominate, because of the impedance matching effects of the transfer function of the closed, cylindrical pipe. The figure shows the frequencies of the harmonics of the sound and those of the impedance maxima in the cylindrical pipe (titled "resonances"), which are approximately at nf_1 where n is an odd integer and f_1 is the frequency of the lowest resonance. (Because of frequency-dependent end effects, these frequencies are not exactly harmonic.) The effects of resonances on harmonics are discussed in more detail later. This spectrum is included to allow comparison of the harmonics measured simultaneously inside and outside the mouth. Above about 1.5 kHz, the spectrum has an increased broadband component. This is the (filtered) sound of the injected acoustic current. Some of the sound injected into the mouth is radiated through the opening lips and the yidaki. Some also leaks through the cladding of the current source directly into the radiation field. Because of this unavoidable contamination of the radiated yidaki sound by the acoustic current used to measure the impedance, all sound spectra shown in subsequent figures were measured immediately following the impedance measurement, during the same, sustained playing gesture. This is also the case in Fig. 4(c).

Figure 4(d) shows the acoustic impedance of the vocal tract during playing. The acoustic impedance was derived from the signal recorded inside the mouth during the injected sound as described above. Frequencies below 200 Hz (the lowest frequency in the injected current) are omitted. To remove the very large signal produced by the vibrating lips, five points centered on each harmonic of the lip frequency up to the 16th have been removed and replaced with a linear interpolation. The resulting data have been smoothed by a linear average over a window of 53.8 Hz and are presented on a linear rather than a logarithmic scale. This process is used hereafter to show Z_{tract} when measured during playing.

Comparing Fig. 4(b) or 4(c) with Fig. 4(d) shows that the peaks in the vocal tract impedance occur at frequencies at which the spectral envelope of the radiated sound has minima. This is considered further, below.

When considering the spectral envelope of Fig. 4(b) or 4(c), one should remember that human hearing sensitivity declines rapidly below about 300 Hz. Consequently, despite their relatively large amplitude, the fundamental and lower

harmonics are not very loud. Further, they vary little during playing. The formants, on the other hand, occur at frequencies in the range of maximum sensitivity of the ear, and they change in response to changes in mouth configuration. It is these formants and the variation in them that contribute most of the interest in yidaki performance.

B. Relationship between the output sound and the impedance of the player's vocal tract

The playing frequency is close to but slightly above that of the first resonance of the pipe, in accordance with Fletcher's (1993) analysis of an "outward swinging door" valve that opens under excess pressure on the upstream side and closes under excess pressure on the downstream side [notated (+, -)]. The signal radiated by this cylindrical yidaki has stronger odd harmonics, especially for low frequencies, where these harmonics fall close to the resonances of the instrument, which are indicated by vertical lines in Fig. 4(c). The broadband component of the radiated signal (largely due to the spectrometer signal leaking through the player's lips into the yidaki) is visible, especially at high frequencies. This broadband spectrum has an envelope that resembles the inverse of the vocal tract impedance, which is discussed below.

The signal in the mouth in Fig. 4(a), which is due to the interaction of the flow through the vibrating lips with the vocal tract, shows no strong difference between even and odd harmonics, because the resonances in the tract are much broader than the frequency differences between the harmonics of the lip vibration.

The vocal tract impedance shows broad peaks at approximately 1.5, 2.1, and 2.8 kHz. A weak peak below 500 Hz is often seen when players mime playing [data not shown, but see Fletcher *et al.* (2001)], but here it is not seen: it is possibly obscured by the strong signal from the vibrating lips.

Comparing the vocal tract impedance spectrum Z_{tract} in Fig. 4(d) with the radiated sound spectrum in Fig. 4(c) one notes that, when Z_{tract} is sufficiently large, the envelope of the radiated sound spectrum is low. This correlation was evident in many such spectra, both from cylindrical pipes and flared yidakis (data not shown). Figure 5 shows the results of independent measurements of the high tongue configuration for the three players described above. The minima and maxima in the spectral envelope of the radiated sound that fell in the range 1.0 to 2.2 kHz were recorded, as were the extrema in the vocal tract impedance measured immediately previously. In Fig. 5, the frequency of each minimum in the spectral envelope of the radiated sound is plotted against that of the nearest maximum in the impedance spectrum (filled symbols) and the frequency of each maximum in the sound envelope is plotted against that of the nearest minimum in the impedance spectrum. The correlation is excellent (the slope is 0.93 and the correlation coefficient is 0.98). The impedance maxima correspond almost exactly to minima in the spectral envelope of the sound, while maxima in the sound spectrum occur on average at frequencies slightly above those of the minimum in the impedance.

Why does a peak in Z_{tract} reduce the level of the radiated sound? The connection is a little obscure and requires careful

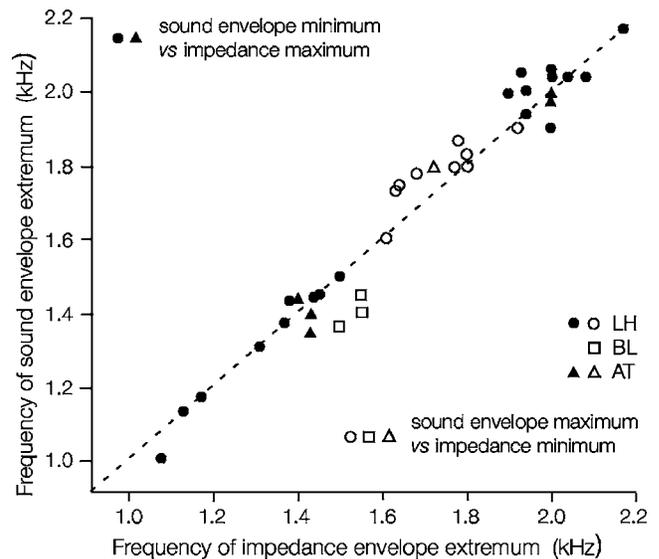


FIG. 5. On this graph, each filled symbol plots the frequency of a minimum in the spectral envelope of the radiated sound against the frequency of the nearest maximum in the impedance spectrum of the vocal tract. Each open symbol plots the frequency of a maximum in the spectral envelope of the radiated sound against the frequency of the nearest minimum in the measured impedance spectrum. No clear impedance maxima were evident for player BL in the measured range. The dashed line is the line of equality.

analysis. For this reason it will be discussed only briefly here, but is treated in detail in our companion theory paper (Fletcher *et al.*, 2006). In most lip-valve instruments, the maxima in Z_{instr} are very much larger than those in Z_{tract} . For the yidaki, in the frequency range of interest (1 to 3 kHz), this is not the case, for two or three reasons. First, the yidaki has a larger cross section than does the mouth with the tongue raised and so it has a relatively small characteristic impedance. Second, in a traditional yidaki, wall losses due to roughness in the bore of a genuine instrument may also be important, though not for the PVC pipe used here. In Fig. 4(d), the peak in Z_{tract} at 1.5 kHz has a value of about 8 MPa s m^{-3} for this configuration when the tongue is raised. The impedance of the pipe has peaks in this frequency range of 3 to 10 MPa s m^{-3} . Consequently, these broad peaks in Z_{tract} give rise to a minimum in the acoustic flow U in the yidaki, at the lips. For any given value of the transfer function between the two ends of the yidaki, a small U at the input yields a small acoustic pressure at the output. Of course, the transfer function of a pipe is a strong function of frequency and has maxima at approximately $f_{\text{pipe}} = (2n + 1)c/4L$, where c is the speed of sound, L is the length of the pipe, and n is an integer. (At these frequencies, the pipe is a good impedance transformer to match the relatively low radiation impedance.) But the resonances of the pipe are closely spaced in frequency compared to those of the tract, so that several harmonics of the played sound will fall within a formant produced by a resonance of the vocal tract. This also explains the shape of the broadband component of the radiated spectra, discussed above.

Figure 6 presents another example of a measurement with the high tongue configuration. We propose that the impedance maxima in Z_{tract} measured just inside the lips shown in Figs. 4(d) and 6(c) are due to resonances of the upper

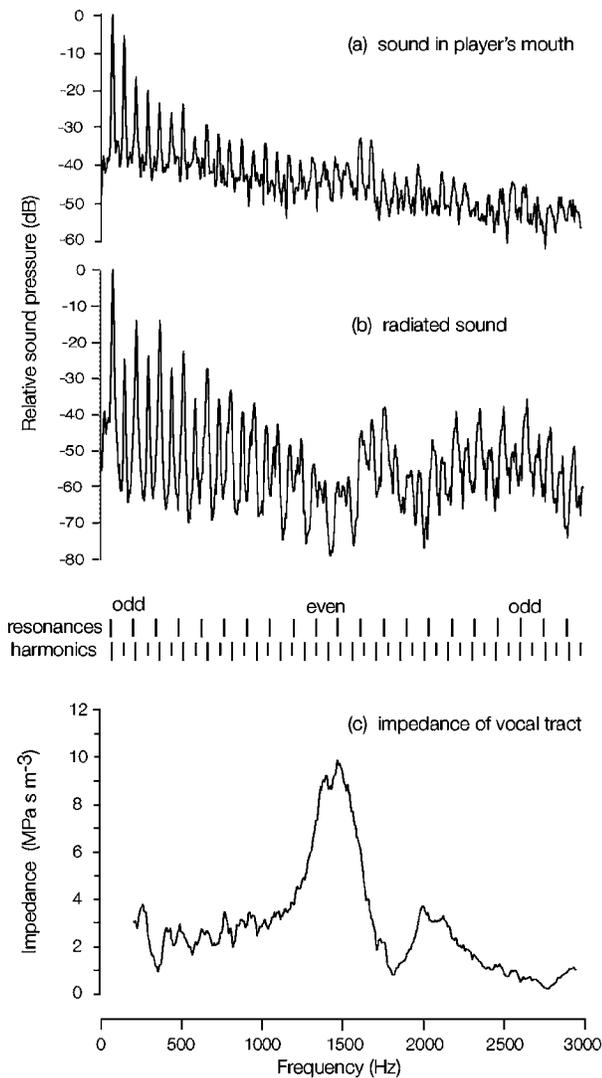


FIG. 6. A recording in the high-tongue configuration, but also illustrating the effects of resonance coincidence. The top graph is the sound spectrum measured inside the mouth, the middle is the sound spectrum outside the end of the yidaki at the same time, and the bottom is the impedance spectrum inside the mouth during playing. At low and high frequencies, the odd harmonics coincide with resonances of the pipe (vertical dashes). However, at around 1.5 kHz, even harmonics coincide with the resonances.

vocal tract (i.e., the airway between the lips and the glottis). Mukai (1989) reports that experienced wind players perform with the glottis almost closed. A nearly closed glottis produces relatively strong resonances at high frequencies because the coefficient of reflection is large. Consequently, at frequencies above several hundred Hz, the impedance seen at the lips is approximately that of an irregular tube closed at the glottis. The lungs, on the other hand, would produce a termination that is essentially resistive at high frequencies, and therefore an airway with open glottis exhibits rather weak resonances (data not shown). We propose that experienced yidaki players, like other wind players, also perform with the glottis nearly closed and that this is necessary to produce the relatively strong resonances that give rise to strong formants in the output sound, as discussed below. In a number of styles of yidaki playing, the glottis is used as an

additional vibrating signal source, so players will presumably be used to keeping the glottis in a nearly closed configuration.

The resonances of the tract in playing are therefore somewhat analogous to those used to produce speech, the differences being that for speech the glottis is the vibrating signal source rather than the lips, and the lips are open. Vocal formants in the kHz range occur at frequencies which produce standing waves in the tract with pressure antinodes near the glottis and a flow antinode at the lip opening (Sundberg, 1977). In other words, they occur when the vocal tract is a most effective impedance matcher between the high impedance at the glottis and the low impedance of the radiation field outside the mouth. The formants radiated by the yidaki also occur when there is an impedance minimum at the lips and, we hypothesize, a pressure antinode near the glottis when it is nearly closed. Of course, in the sustained vowels of speech, the lips are at least somewhat open, whereas in yidaki playing they are almost shut. In speech, the mouth opening affects primarily the first vocal formant (F_1 , which occurs below about 1 kHz for all vowels). It is the second formant (F_2) that is of interest here, because it falls approximately in the range 1 to 2 kHz. The frequency of the F_2 in speech depends somewhat on mouth opening. Consequently, the strong yidaki formant could be expected to occur at frequencies comparable with, but not equal to, those of second speech formants for a similar tract configuration. Thus playing with a mouth configuration similar to that required to produce the vowel /i/, for example, will produce a yidaki sound whose formant frequency is similar to but not necessarily equal to that of the second formant in the vowel /i/. Further, to produce a yidaki formant, the mouth configuration must provide peaks in the tract impedance that have sufficiently high amplitude, so some vowel shapes with low tongue may not produce a clear formant, as is discussed below.

Figure 6 illustrates another effect that influences formants in the output sound. Again, the playing frequency is slightly above that of the first resonance—about 3 Hz in this case. Consequently, at a frequency approaching 1 kHz, the odd-even difference in the sound spectrum disappears, because in this frequency range the harmonics fall almost midway between resonances. At around 1.5 kHz, on the other hand, it is the even harmonics that benefit from the resonances of the pipe. This range is not far below the formant due to the vocal tract resonance. It might be possible, in principle, to observe broad and weak formants in the sound spectrum due only to this effect of an even or an odd harmonic happening to fall on a resonance of the yidaki: we call this the “harmonic coincidence” effect. For an experienced yidaki player, formants produced by this effect would be relatively small compared to those produced by the minima in the spectral envelope that coincide with peaks in Z_{tract} . Consequently, in this figure, as in Fig. 4, the maxima in the impedance of the vocal tract coincide with minima in the spectral envelope of the radiated sound. However, in Fig. 6 (but not Fig. 4), the formants at about 1.6 kHz is somewhat assisted by the near coincidence of a harmonic of the lip motion (here an *even* harmonic) with a resonance of the in-

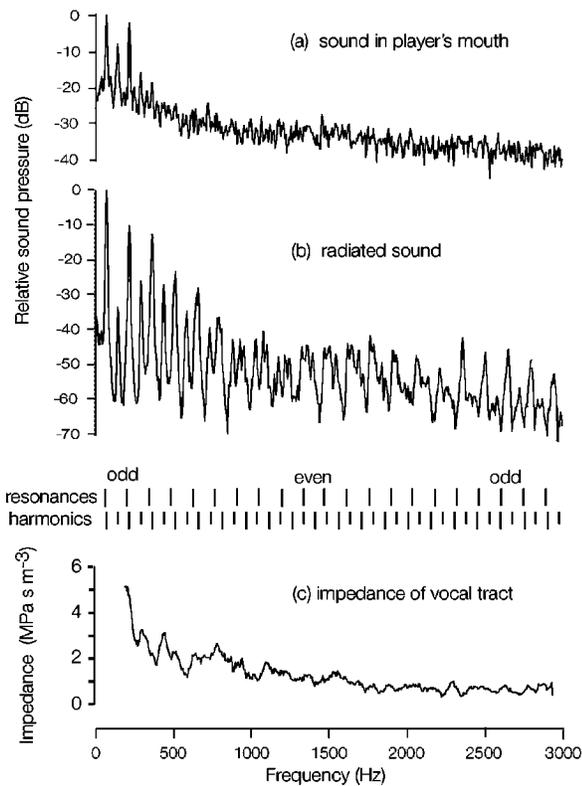


FIG. 7. The graphs are as for Figs. 4 and 6, but these data were taken during a note played with the low tongue configuration.

strument. As the player can readily make small adjustments to the playing frequency, it is possible that the playing frequency is sometimes adjusted to take advantage of harmonic coincidence, though this was not investigated here.

Figure 7 shows the impedance measured inside the mouth and the sound produced for a configuration in which the player held the tongue low in the mouth. The peaks in Z_{tract} were not as large as for the high tongue configuration, and consequently there is none of the shaping of the spectral envelope of the sound as is produced with the high tongue figuration. The impedance of the vocal tract is so low over most of the frequency range that only the strong low harmonics of the lip motion produce components that are clearly visible above the turbulent noise present inside the mouth. There is no clear formant comparable to those near 1.7 kHz in Figs. 4 and 6. The coincidence effect does increase the odd harmonics around 2.5 kHz. In this example, and in others measured in the low tongue configuration, the odd harmonics in the signal recorded in the mouth were stronger than the neighboring even harmonics. Whether this is due to the shape of the lip opening (discussed below) or to sound transmission from the bore of the yidaki into the relatively low impedance load in the mouth, or to another cause, we do not know. However, one of its consequences might be the absence of a clear coincidence effect around 1.5 kHz.

C. Motion of the lips and air jet during playing

Images of the playing lips were produced in two different ways. In one, a high speed camera operating at 1000 frames per second filmed the lips from both the front and the

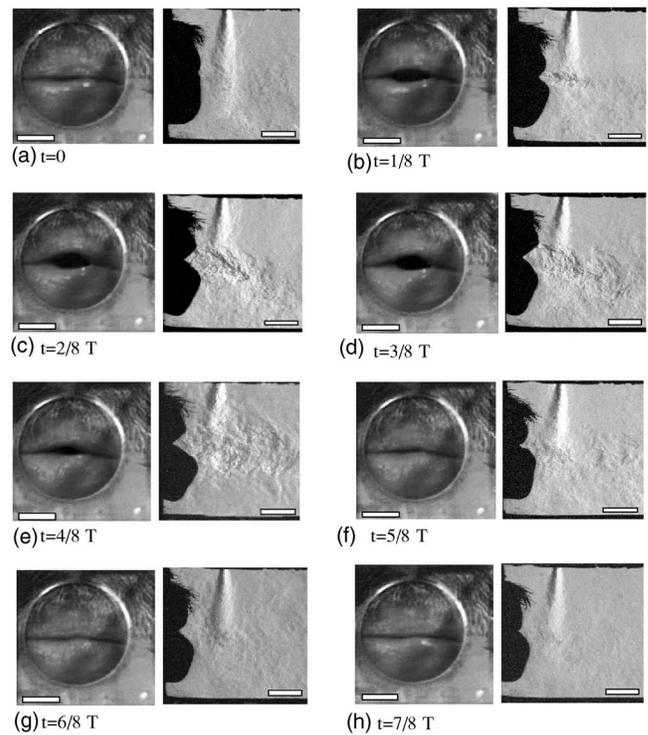


FIG. 8. The images on the left show one cycle of vibration of the player AT's lips, from the front, photographed with a high speed camera. The bar has length 10 mm. The side views (at right) were obtained stroboscopically by the schlieren flow visualization technique and are matched with the high speed images to show one complete cycle. The scale bars have length 10 mm.

side using the arrangement shown in Fig. 2. The side-view pictures thus taken are not shown here. Instead we present in side view the schlieren images, which show both the air jet and the lip position. These were produced stroboscopically using the arrangement shown in Fig. 3. Images from these two series are shown in Fig. 8. The images on the left of each pair are sequential pictures from one cycle. The pictures on the right, however, were each obtained in different cycles from the stroboscopic series and were matched to those from the high speed series by matching the shape of the lips at opening. The high speed film (both front and side view) is on our web site (Music Acoustics, 2005). The lip shapes shown here are qualitatively similar to the results reported by Wiggins (1988).

Figure 9 shows data from sets of images such as those on the left in Fig. 8, and the side view images taken simultaneously via the mirror (pictures not shown). The data are averaged over ten cycles from a series made with the high tongue configuration. The maximum camera speed (1000 frames per second) limits the time resolution, and hence limits the maximum frequency in an experimental spectrum $A(f)$ to 500 Hz. However, the shape of $A(t)$ is used as an input in a numerical model presented in our companion theory paper (Fletcher *et al.*, 2006). Yoshikawa (1995) reported the motion of the lips of horn players. In the low end of the range, their lips moved mainly along a horizontal axis (i.e., parallel to the bore of the mouthpiece), with a smaller opening motion in the vertical direction. Like that of the horn players in the low register, the lip motion reported here thus conforms

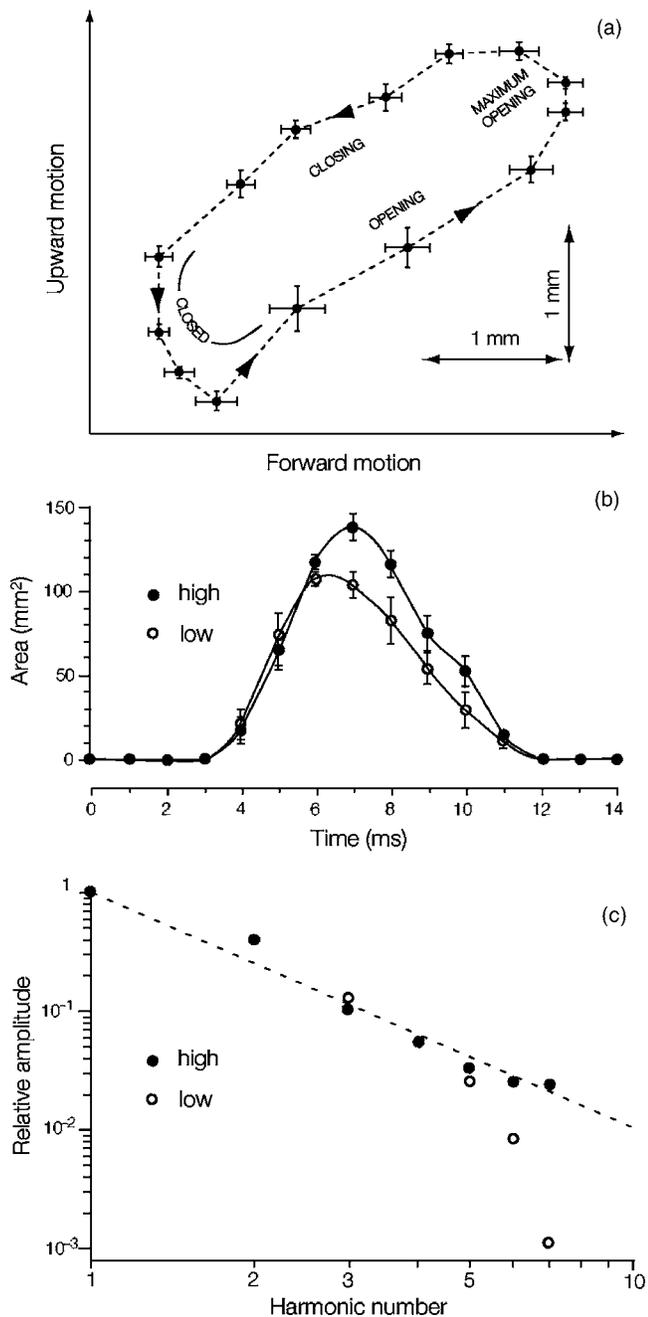


FIG. 9. Data from images such as those in Fig. 8, averaged over ten cycles. (a) shows how the average position of the upper lip varies throughout a complete cycle (bars show the standard error). (b) shows how the average area of the open space between the lips varies with time throughout a cycle for the high and low tongue configurations. (c) shows the relative amplitude of the harmonic components of the periodic variations in the lip area. The dashed line indicates the situation in which the harmonic components vary as n^{-2} .

to the “outward swinging door model” used to explain the operation of lip-valve instruments (Fletcher and Rossing, 1998) and the yidaki (Hollenberg, 2000).

The schlieren images in Fig. 8 show the motion of the jet emitted from the lips by mixing the jet with CO₂ from the “curtain” introduced as shown in Fig. 3. The density of carbon dioxide is about 1.5 times greater than that of air. However, the total volume of CO₂ is small so there is little reason to suspect that this alters significantly the behavior of the

yidaki in general. Indeed, the period of the sound pressure oscillation (14.3 ms) remained substantially unchanged when the CO₂ was supplied to form the curtain. To reduce the effect of the CO₂ jets on the air jet, the flow of CO₂ was regulated so that the CO₂ curtain becomes almost invisible at the level of the opening of the lips. Some aspects of the motion of the jet are explained by the mouth opening. In Fig. 8(b), the lips have just opened and the jet is well-defined and narrow. In Figs. 8(b)–8(e) inclusive, the lips remain open and the jet grows broader. This is possibly the result of changing geometry of the jet separation from the lips as they open. It is interesting that the jet deviates noticeably downward from the axis of the yidaki [Figs. 8(e) and 8(f)]. The downward deviation may be caused by the changing geometry of the player’s lips as they open or it may be due to a momentum transfer between the jet and the descending CO₂ stream. Once the lips close [Fig. 8(f)] the momentum of the jet keeps it moving and the disturbance of the “curtain” gradually disappears.

In a widely used Poiseuille-flow model for wind instruments, the relation between the flow of air into the instrument and pressure difference between the mouth and the instrument is calculated assuming conservation of energy between the mouth and the jet, followed by dissipation of kinetic energy in the jet when it mixes with the air in the instrument. When the air jet emerges from the lips it experiences viscous drag from the surrounding air, which leads to turbulent mixing because of the high speed of the jet. This can be clearly seen in Fig. 8. The important thing for sound generation in the instrument, however, is simply the input volume flow at a particular frequency multiplied by the instrument impedance at that frequency.

D. Idiomatic effects: Circular breathing and vocalization

Figure 10 shows aspects of the sound produced for another important playing configuration: the inhalation phase necessary for “circular breathing,” during which the performer plays using the air in the inflated cheeks, with his mouth sealed from the lower vocal tract with the soft palate. This radical change in the vocal tract geometry makes a substantial change in the timbre and the necessity of frequent inhalation means that the change occurs often in continuous playing. Idiomatic playing makes a virtue of this necessity, so that the sound produced using the cheeks as reservoir is made an integral element of the rhythmic structure. Many of the rhythms used require simple alternation of inhalation and exhalation through the instrument. Typical mouth pressures vary from 0.8 to 2 kPa and flow rates from 0.1 to 0.3 l/s. The volume of air that can be expelled from the cheeks supports loud playing for only a fraction of a second, during which most players are unable to inspire deeply. Consequently, a simple alternation of inhalation and normal playing allows only shallow breathing.

For this reason, players will sometimes play for several seconds using the air in the lungs, and then refill the tidal volume during a series of three or four alternations between brief inhalations (while playing with the cheek reservoir) and

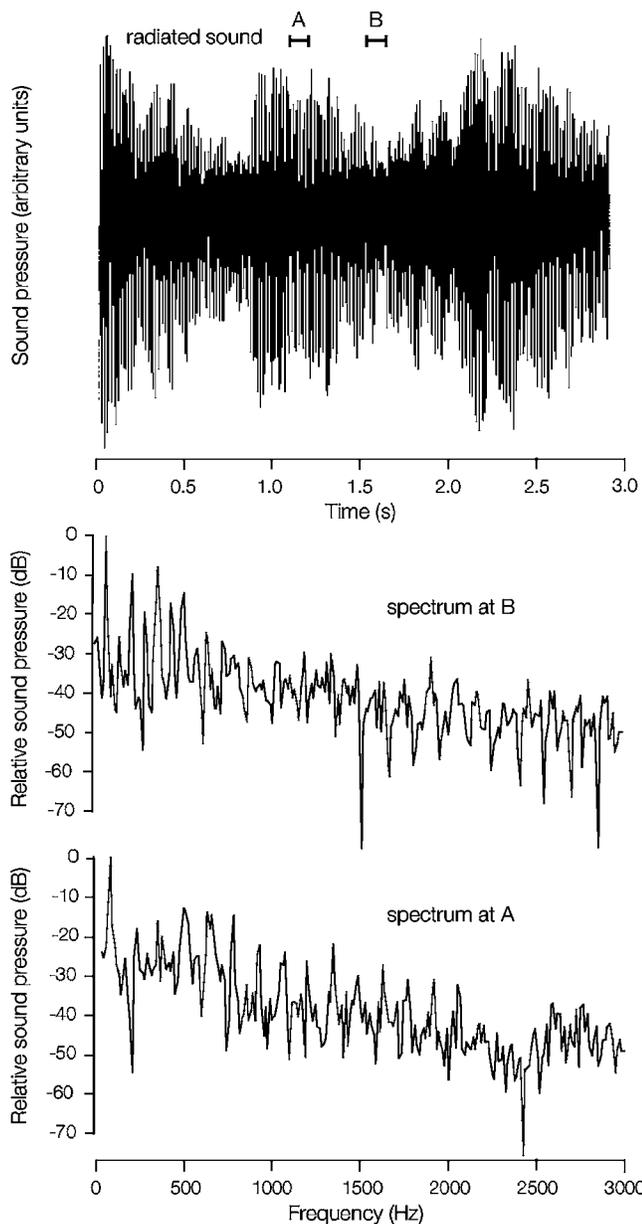


FIG. 10. Top: an oscillogram of a sound sample during which the player inhales through the nose three times, while continuing to play—an example of one of the common rhythms used in circular breathing. The amplitude falls during the inhalations and is largest during normal playing. Spectra of sounds during normal playing (a) and inhalation (b) are shown.

normal playing. This may be used when preparing for or recovering from the use of a high flow rate, such as may be used to play one of the upper resonances.

An example of a regular series of inhalations is shown in Fig. 10. The waveform shows a series of three exhalations (large amplitude), each followed by a brief inhalation (small amplitude). The amplitude decreases during each of the inhalations, which may be explained by the application of the Young-Laplace relation to the muscles of the cheeks: it is easier to apply a higher pressure with a given muscular tension when the cheeks are highly curved. The finite mouth volume and the consequent brief period during inhalation (typically 0.2 s for this player) had the unfortunate conse-

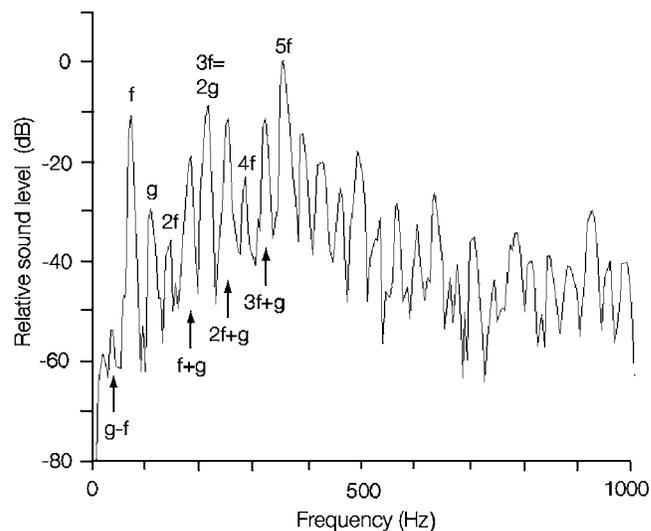


FIG. 11. The spectrum of a sound produced by playing a note with frequency $f=70$ Hz and simultaneously vocalizing at a frequency g , where $g=3f/2=105$ Hz.

quence for this study that it proved impossible to make acceptable impedance measurements of the tract during this gesture.

The spectra in Fig. 10 contrast the spectra of the sound produced during one of the inhalations, during which the mouth alone was the resonator upstream from the lips and during the intervening exhalation.

A further series of different timbres can be obtained by vocalizing and playing at the same time. In this case, both the vocal folds and the lips act as pressure-controlled valves, operating at opposite ends of the vocal tract. Due to the strongly nonlinear interaction among lip motion, vocal fold motion, and the air flow, a range of heterodyne components are generated. Simply, if the vocal folds have opening area $S_1 = \sum a(n) \sin(n\omega_1 t)$ and the lips have opening area $S_2 = \sum b(n) \sin(n\omega_2 t)$, then the total flow, ignoring tract impedance effects, is proportional to $S_1 S_2$ and thus has components at all frequencies $n\omega_1 \pm m\omega_2$. An example is shown in Fig. 11. Player LH plays a steady note at frequency $f=70$ Hz, while simultaneously vocalizing at a frequency $g=3f/2=105$ Hz, a musical fifth above. The first several harmonics and heterodyne components are indicated in the figure. Because of the harmonic relation between the two frequencies, these components are equally spaced at frequencies $f/2 = g/3 = 35$ Hz. Consequently, the pitch is heard at one octave below the lip fundamental, as the sound files demonstrate (Music Acoustics, 2005).

IV. CONCLUSIONS

When the player's tongue is raised close to the hard palate, the acoustical impedance spectrum of the vocal tract has maxima with values of 2–10 MPa s m⁻³ in the range 1 to 3 kHz—comparable with or larger than a typical impedance maximum of the yidaki in that frequency range. At the frequencies of these maxima, the spectral envelope of the radiated sound has minima, because the frequency components of acoustic current in the air jet entering the instrument

are reduced. In the frequency bands lying between the impedance maxima, the lower impedance of the tract produces a stronger acoustic current through the lips, resulting in a characteristic, strong formant in the radiated sound. These formants in the radiated sound correspond approximately to those of the second formant frequency produced in speech with the same tract configuration. The large variation between impedance maxima and minima, the variation that produces strong formants in the radiated sound, is consistent with a glottis nearly closed, the configuration used by experienced wind players studied by Mukai (1989, 1992). This suggests that learning to play with the glottis nearly closed may be important in developing technique on the yidaki. Broad and weak formants can also be observed when groups of even or odd harmonics coincide with a bore resonance.

The lips are open for about half of each cycle and operate approximately as “outward swinging doors,” as described by Yoshikawa (1995) for the low range of the horn. The motion of the jet is relatively simple and supports the approximation of one-dimensional motion made in several analyses (Fletcher, 1993; Hollenberg, 2000). The present paper has been concerned entirely with experimental measurements. A theoretical analysis and justification of some of the conclusions will be presented in a companion paper (Fletcher *et al.*, 2006).

ACKNOWLEDGMENTS

We thank John Tann for technical assistance and the Australian Research Council for support. BL worked on this project during a vacation scholarship.

- Amir, N. (2004). “Some insights into the acoustics of the didjeridu,” *Appl. Acoust.* **65**, 1181–1196.
- Amir, N., and Alon, Y. (2001). “A study of the didjeridu: normal modes and playing frequencies,” *Proc. International Symposium on Musical Acoustics*, Perugia, edited by D. Bonsi, D. Gonzalez, and D. Stanzial, pp. 95–98.
- Backus, J. (1985). “The effect of the player’s vocal tract on woodwind instrument tone,” *J. Acoust. Soc. Am.* **78**, 17–20.
- Berio, L. (1966). *Sequenza V; Solo Trombone* (Universal, New York).
- Brass, D., and Locke, A. (1997). “The effect of the evanescent wave upon acoustic measurements in the human ear canal,” *J. Acoust. Soc. Am.* **101**, 2164–2175.
- Caussé, R., Goepf, B., and Sluchin, B. (2004). “An investigation on ‘tonal’ and ‘playability’ qualities of eight didgeridoos, perceived by players,” *Proc. International Symposium on Musical Acoustics*, Nara, Japan.
- Clinch, P. G., Troup, G. J., and Harris, L. (1982). “The importance of vocal tract resonance in clarinet and saxophone performance—a preliminary account,” *Acustica* **50**, 280–284.
- Elliott, J. S., and Bowsher, J. M. (1982). “Regeneration in brass wind instruments,” *J. Sound Vib.* **83**, 181–217.
- Epps, J., Smith, J. R., and Wolfe, J. (1997). “A novel instrument to measure acoustic resonances of the vocal tract during speech,” *Meas. Sci. Technol.* **8**, 1112–1121.
- Erickson, R. (1969). *General speech: for trombone solo (with theatrical effects)* (Smith, Baltimore).
- Fletcher, N., Hollenberg, L., Smith, J., and Wolfe, J. (2001). “The didjeridu and the vocal tract,” *Proc. International Symposium on Musical Acoustics*, Perugia, edited by D. Bonsi, D. Gonzalez, and D. Stanzial, pp. 87–90.
- Fletcher, N. H. (1983). “Acoustics of the Australian didjeridu,” *Australian Aboriginal Studies* **1**, 28–37.
- Fletcher, N. H. (1993). “Autonomous vibration of simple pressure-controlled valves in gas flows,” *J. Acoust. Soc. Am.* **93**, 2172–2180.
- Fletcher, N. H. (1996). “The didjeridu (didgeridoo),” *Acoust. Aust.* **24**, 11–15.
- Fletcher, N. H., and Rossing, T. D. (1998). *The Physics of Musical Instruments*, 2nd ed. (Springer-Verlag, New York).
- Fletcher, N. H., Smith, J., Tarnopolsky, A., and Wolfe, J. (2005). “Acoustic impedance measurements—correction for probe geometry mismatch,” *J. Acoust. Soc. Am.* **117**, 2889–2895.
- Fletcher, N. H., Hollenberg, L. C. L., Smith, J., Tarnopolsky, A. Z., and Wolfe, J. (2006). “Vocal tract resonances and the sound of the Australian didjeridu (yidaki) II. Theory,” *J. Acoust. Soc. Am.* **119**, 1205–1213.
- Hollenberg, L. (2000). “The didjeridu: Lip motion and low frequency harmonic generation,” *Aust. J. Phys.* **53**, 835–850.
- Mukai, M. S. (1992). “Laryngeal movement while playing wind instruments,” in *Proc. International Symposium on Musical Acoustics*, Tokyo, Japan, pp. 239–242.
- Mukai, S. (1989). “Laryngeal movement during wind instrument play,” *J. Otolaryngol. Jpn.* **92**, 260–270.
- Music Acoustics (2005). <http://www.phys.unsw.edu.au/~jw/yidakididjeridu.html>
- Smith, J. R., Henrich, N., and Wolfe, J. (1997). “The acoustic impedance of the Bøehm flute: standard and some non-standard fingerings,” *Proc. Inst. Acoust.* **19**, 315–320.
- Sundberg, J. (1977). “The acoustics of the singing voice,” *Sci. Am.* **236**, 82–91.
- Tarnopolsky, A., Fletcher, N., Hollenberg, L., Lange, B., Smith, J., and Wolfe, J. (2005). “The vocal tract and the sound of a didgeridoo,” *Nature (London)* **436**, 39.
- Tarnopolsky, A. Z., and Fletcher, N. H. (2004). “Schlieren flow visualisation technique: applications to musical instruments, especially the didjeridu,” *Proc. 18th International Congress in Acoustics*, Kyoto, Japan, 4–9 April.
- Tarnopolsky, A. Z., Lai, J. C. S., and Fletcher, N. H. (2000). “Flow structures generated by pressure-controlled self-oscillating reed valves,” *J. Sound Vib.* **247**(2), 213–226.
- Wiggins, G. C. (1988). “The physics of the didgeridoo,” *Phys. Bull.* **39**, 266–267.
- Wolfe, J., Smith, J., Tann, J., and Fletcher, N. H. (2001). “Acoustic impedance of classical and modern flutes,” *J. Sound Vib.* **243**, 127–144.
- Wolfe, J., Tarnopolsky, A. Z., Fletcher, N. H., Hollenberg, L. C. L., and Smith, J. (2003). “Some effects of the player’s vocal tract and tongue on wind instrument sound,” *Proc. Stockholm Music Acoustics Conference (SMAC 03)*, edited by R. Bresin, Stockholm, Sweden, pp. 307–310.
- Yoshikawa, S. (1995). “Acoustical behavior of brass player’s lips,” *J. Acoust. Soc. Am.* **97**, 1929–1939.