

Vocal tract resonances and the sound of the Australian didjeridu (yidaki) II. Theory

N. H. Fletcher^{a)}

*Research School of Physical Sciences and Engineering, Australian National University,
Canberra 0200, Australia*

L. C. L. Hollenberg

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

J. Smith, A. Z. Tarnopolsky, and J. Wolfe

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

(Received 7 August 2005; accepted 8 November 2005)

The didjeridu (didgeridoo) or yidaki of the Australian Aboriginal people consists of the narrow trunk of a small Eucalypt tree that has been hollowed out by the action of termites, cut to a length of about 1.5 m, smoothed, and decorated. It is lip-blown like a trumpet and produces a simple drone in the frequency range 55 to 80 Hz. Interest arises from the fact that a skilled player can make a very wide variety of sounds with formants rather like those of human vowels, and can also produce additional complex sounds by adding vocalization. An outline is given of the way in which the whole system can be analyzed using the harmonic-balance technique, but a simpler approach with lip motion assumed shows easily that upper harmonics of the drone with frequencies lying close to impedance maxima of the vocal tract are suppressed, so that formant bands appear near impedance minima of the vocal tract. This agrees with experimental findings. Simultaneous vibration of the player's lips and vocal folds is shown to generate multiple sum and difference tones, and can be used to produce subharmonics of the drone. A brief discussion is given of player preference of particular bore profiles. © 2006 Acoustical Society of America. [DOI: 10.1121/1.2146090]

PACS number(s): 43.75.Fg, 43.70.Aj, 43.75.Rs, 43.72.Ct [DD]

Pages: 1205–1213

I. INTRODUCTION

The Aboriginal peoples of Australia have lived in comparative isolation in this continent for something in excess of 40 000 years. They are divided into many nomadic tribes, each living in traditional lands and each speaking a different language but with many common customs. These people have developed several devices that have attracted wide attention, the best known being the boomerang, which returns to the hands of the thrower after quite a long circular flight, and the didjeridu, usually spelled “didgeridoo” in nonacademic literature. They have also developed a version of the “bullroarer” that is similar to that found in many other ancient cultures, and discovered how to imitate bird songs and other sounds using a “gum-leaf” from a Eucalypt tree pressed against the lips.¹

Lip-blown wind instruments are common in many ancient cultures, their form depending upon availability of simple tube-like or conical structures to define the air column. Conch shells, for example, have been popular, as also have been tubes derived from the hollow stems of bamboo. The didjeridu was developed by the Yolngu tribes whose traditional lands are in what is now known as Arnhem Land, to the east of the modern city of Darwin on the central northern coastline of Australia. In the tribal language the instrument is called a yidaki or yiraki. In comparatively recent

times its manufacture has spread to most Aboriginal craft communities and thence to many modern makers in other countries. Quite apart from its traditional use for the accompaniment of songs and ceremonies,^{2,3} the didjeridu has also become widely used in “contemporary” music bands.⁴

The didjeridu is a simple wooden tube, typically about 1.5 m in length but sometimes made shorter, particularly in Central Australia. The tube is derived from the action of termites, which are common in Arnhem Land and which eat out the dry wood at the core of living Eucalypts, which is the major tree genus in Australia. Some species have quite narrow trunks, and the extent of termite damage can be assessed by knocking on the trunk with a heavy stick. When an adequately eaten trunk is found, the tree is cut down, traditionally using a sharpened stone, and a suitable section is cut from near its base. The termites leave a honeycomb-like gal-leried structure that can be cleaned out with a long stick and the inside of the tube rubbed to moderate smoothness, a process that is possible because of the growth rings in the wood. Some work with a sharp stone may be necessary at the two ends to give a smooth termination. The bark is stripped from the outside of the tube and it is then smoothed and usually painted with traditional tribal or family totem designs, using brown, black, and white pigments made from soil, charcoal, and minerals held together with plant gums. Finally, the smaller blowing end of the tube is smoothed for the player's lips by application of a ring of beeswax.

The sound of the didjeridu is a low-pitched drone with frequency usually about 70 Hz, maintained for minutes at a

^{a)}Also at School of Physics, University of New South Wales, Sydney 2052, Australia. Electronic mail: neville.fletcher@anu.edu.au

time by the technique of “circular breathing” in which the cheeks are blown out to hold enough air to maintain sound production, the soft palate is momentarily closed, and there is a quick intake of breath through the nose. This technique is not unique to the didjeridu but has been used by traditional Indonesian flute players, and is now common among professional orchestral players of some woodwind instruments. The drone sound is normally modified by introducing emphasized frequency bands, known as formants by analogy with human vowel sound spectra, through change in the configuration of the upper vocal tract. These formants are usually introduced in a rhythmic manner to produce a musical pattern. Occasional accents are also introduced by momentary leaps to the second mode of the air column, at a frequency between about 2.5 and 3 times the drone frequency, depending upon the shape of the instrument bore. Finally, while maintaining the lip vibration, the player may also occasionally vibrate his vocal folds as in singing to produce additional sounds or even subharmonics of the drone. These techniques are often used to imitate the sounds made by native animals or to illustrate the text of a song or story.

The acoustics of the didjeridu tube is essentially trivial, since it consists of a tube of defined though irregular geometry with enhanced wall losses because of roughness and porosity. Details of these parameters determine the distinction between an instrument considered “good” by the player and one regarded as “poor,”⁵⁻⁷ as will be discussed later. Interest however centers upon the coupling of the vocal tract to the instrument and on the vibration of the player’s lips. There has been rather little investigation of these problems, except for work by Fletcher,^{8,9} Wiggins,¹⁰ Hollenberg,¹¹ and members of the present project team.¹²

The purpose of the present paper is to give a formal analysis of the acoustics of the played didjeridu, including the important contributions of the player’s lungs and vocal tract. Some consideration will also be given to the effect of vocalizing while playing. Only brief mention will be given here of experimental studies by our project team in order to provide a comparison with calculated results. Detailed discussion of the measurements and experiments conducted is given in a companion paper,¹³ and a brief discussion of related experimental results has recently been published.¹⁴

II. ANALOG NETWORK MODEL

Analysis of the acoustic behavior of the entire system consisting of lungs, glottis, vocal tract, lips, and didjeridu tube is most conveniently carried out using an electrical network analog,¹⁵⁻¹⁷ in which voltage represents acoustic pressure and current represents acoustic volume flow. Since the whole system is very complicated, it is helpful to treat it in sections, beginning with the lungs.

A. Lung impedance

The lung is a complex quasifractal structure of branching tubules.¹⁸⁻²⁰ In a simplified version, starting just below the vocal folds the main tube branches into two bronchi which continue to branch into pairs of tubules through 16 orders. After 16 orders of branching the tubules are termi-

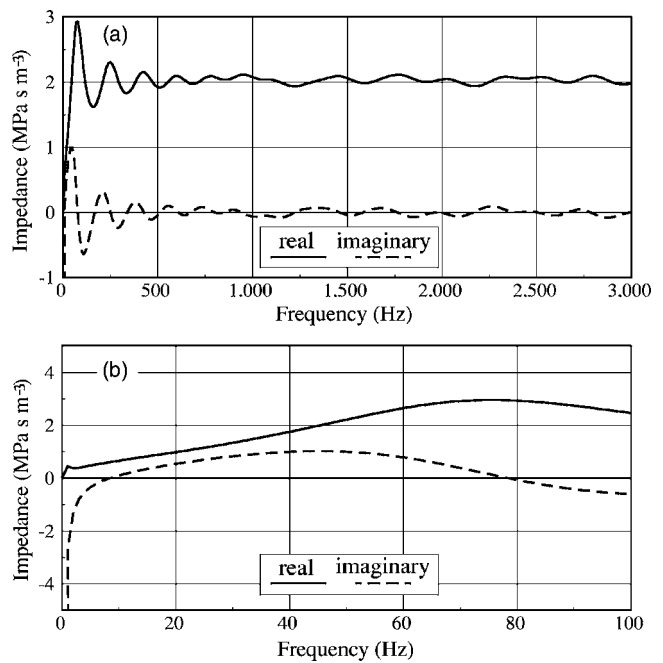


FIG. 1. Calculated input impedance just below the vocal folds for the idealized lung model (a) over an extended frequency range, and (b) over the range of interest for the lip vibrations. The characteristic impedance of the bronchus is about 2 MPa s m^{-3} .

nated by about 65 000 small alveolar sacs which contribute most of the stored air volume. There are, of course, several spatial constraints to be obeyed by this structure so that the tubules and finally the alveolar sacs can fit within the overall volume occupied by the lungs in the body. The whole structure is somewhat elastic, and the pressure inside it is maintained by muscle tension on walls of quite large total mass.

As a simple model for the lungs that is a good approximation to reality, we take the tubule diameter ratio after branching to be 0.8, rather than the idealized value of $2^{-1/2}$, and take the tubule length ratio also to be 0.8. This allows the geometrical packing constraints to be satisfied and gives a total tubule volume of about 1 liter, leaving the remaining 3 liters of lung capacity to be provided by the alveolar sacs. It is straightforward to calculate the acoustic input impedance for this idealized model, and the results are shown in Fig. 1. Over the range of interest for the vibration of the lips, near 70 Hz, the impedance is almost purely resistive and about equal to the characteristic impedance of the bronchus, $2 \times 10^6 \text{ Pa s m}^{-3}$, and the same is true in the range from about 0.5 to 3 kHz that is of interest for the vocal-tract resonances leading to sound formants.

As a consequence of this analysis, it is a good approximation to represent the lungs in the network model as shown in Fig. 2(a). Muscular pressure on the lungs is represented by a dc voltage source in series with a very large inductance representing the mass load (inertance) of the associated muscular tissue. The acoustic compliance of the air volume in the alveolar sacs is represented by a simple capacitance, while the impedance of the network of tubules contributes a simple resistance with magnitude about equal to the characteristic impedance of an infinite tube of diameter equal to that of the subglottal bronchus. Such a model reproduces the behavior shown in Fig. 1.

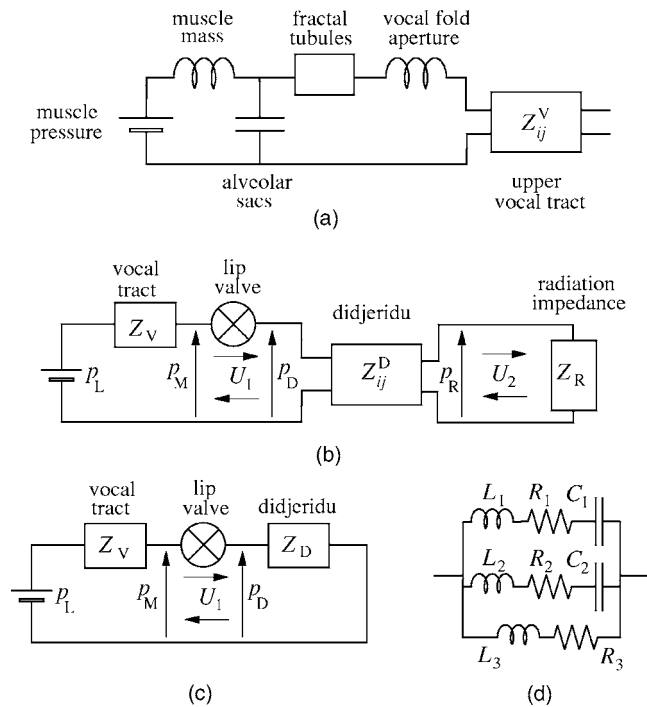


FIG. 2. (a) Electrical network representing the lungs, glottis, and vocal tract of the didjeridu player. (b) Simplified network representing the entire didjeridu+player system. (c) Further simplified network used for calculation. (d) Simple model used for the vocal tract impedance Z_V .

Above the junction of the two bronchi, we find the constriction of the vocal folds. Since these folds are inactive during normal playing, the larynx constriction can be represented in the model by a simple series inductance, as shown. It appears that skilled players on most wind instruments partly close (adduct) their vocal folds so that the area of the remaining aperture is small, a feature first noted and examined in detail by Mukai.^{21,22} This provides a significant series inductance between the upper vocal tract and the lungs and serves to isolate the upper-tract resonances to a large extent from the damping influence of the resistive lung impedance. For typically adducted vocal folds, the opening is about 2×10 mm and the effective fold thickness, including end corrections, perhaps 5 mm, giving an acoustic inductance at frequency ω of about 300ω Pa s m^{-3} , which is about $2\text{--}6$ MPa s m^{-3} in the frequency range 1–3 kHz that is of interest for formants. This must be compared with characteristic impedance $\rho c/S \approx 1$ MPa s m^{-3} of the trachea and the estimated subglottal lung impedance, illustrated in Fig. 1, which is nearly purely resistive and about 2 MPa s m^{-3} in magnitude. It is therefore evident that the vocal folds can have a significant effect on the resonances of the vocal tract provided they are well adducted.

B. Vocal-tract impedance

Above the vocal folds, which have already been discussed, the trachea is a nearly uniform tube of diameter about 20 mm until it reaches the glottis, above which the geometry is complicated by the movable lower jaw and tongue. Magnetic resonance imaging (MRI) studies¹² of one of us (LH) while playing on a modified didjeridu provide

some numerical values for the geometrical parameters for both the “uninflected drone” (low tongue) and “high formant” (high tongue) configurations. As will be seen from the analysis to follow, the acoustical quantity of importance in determining the spectral envelope of the didjeridu sound is the frequency dependence of the input impedance Z_V of the vocal tract at the lips, as shown in Fig. 2(b). This quantity, measured by a modification of an impedance head developed by our group,^{23,24} is discussed in detail in the companion experimental paper.¹³ In the low-tongue position, the mouth simply extends the trachea into a uniform tube up to the lips, giving a rather low value for Z_V . In the high-tongue configuration, normally used to produce sounds with pronounced formants, the tongue is raised along its whole length and brought close to the teeth, so that it forms a narrow passage of effective diameter about 10 mm from the root of the tongue up to the lips. By changing the position of the lower jaw and tongue, it is also possible to convert this passage from almost cylindrical to nearly conical, with obvious influence upon the tract resonances, while the narrow duct diameter near the lips greatly increases the value of Z_V . This effect has been briefly described, both for the didjeridu and for other wind instruments, in a recent conference publication.²⁵

A detailed model for the impedance Z_V of the whole vocal tract at the lips can be built up using the analog network shown in Fig. 2(a), the variations in the geometry of the upper vocal tract being detailed through the coefficients $Z_{ij}^V(\omega)$. In the analysis to follow, however, a very simplified model will be used for the vocal-tract impedance, as shown in Fig. 2(d). This consists of two series-resonant LRC circuits in parallel, representing the resonances of the upper vocal tract, the characteristic impedance $(L/C)^{1/2}$ of these circuits being chosen to match the expected impedance of the vocal tract for the high-tongue configuration, and with a further parallel LR branch representing the impedance of the glottal constriction and the lungs.

C. Didjeridu impedance

The acoustic properties of the didjeridu can be described by four impedance coefficients $Z_{ij}^D(\omega)$, the values of which can be calculated from the measured dimensions of the tube and the estimated roughness of the walls. Most genuine didjeridu tubes flare progressively along their length so that their shape is roughly that of a truncated cone. For such a tube, ideally open at end 2, the impedance maxima at the mouth occur at frequencies ω_n such that

$$k_n L' = n\pi - \arctan\left(\frac{k_n d_1 L'}{d_2 - d_1}\right), \quad (1)$$

where $k_n = \omega_n/c$, c is the speed of sound in air, d_1 is the tube diameter at the blowing end, d_2 is the larger diameter at the open end, L is the tube length, and $L' \approx L + 0.3d_2$ is the effective tube length including the open-end correction.⁹ For a cylindrical tube with $d_2 = d_1$, the mode frequencies ω_n given by (1) follow a sequence 1,3,5,... times $\pi c/2L'$, while for a complete cone of the same length the sequence is 2,4,6,... times $\pi c/2L'$. For an intermediate geometry of small flare, the mode frequencies are given by

$$\omega_n \approx \left(n - \frac{1}{2} \right) \frac{\pi c}{2L} \left\{ 1 + \left[1 + \frac{4(d_2 - d_1)}{\pi^2 d_1 (n - (1/2))^2} \right]^{1/2} \right\}. \quad (2)$$

These frequencies lie between those of the limiting sequences mentioned above for cylindrical and for conical tubes of the same length, and thus might be described as a “stretched quasiharmonic sequence.” A real didjeridu tube, being a natural object hollowed out from a tree trunk by termites, does not have quite this regular form, though the behavior of its lower modes is generally well approximated by the flared-tube relation (2).

Experimental values for the impedance coefficients $Z_{ij}^D(\omega)$ can be obtained by measuring the input impedance of the tube from each end with the opposite end either open or closed, in separate measurements, or can be calculated if the geometry is assumed to be regular. If the shape of the bore of the didjeridu is assumed to be a uniform cylinder of length L , as it may be in some laboratory experiments using plastic pipe, then the impedance coefficients have the simple forms

$$Z_{11}^D = Z_{22}^D = -jZ_0 \cot kL \quad (3)$$

$$Z_{12}^D = Z_{21}^D = -jZ_0 \csc kL,$$

where $Z_0 = 4\rho c / \pi d^2$ is the characteristic impedance of the tube, assumed to have diameter d , and $k \approx \omega/c - j\alpha$. The quantity α allows for viscous and thermal losses to the tube walls according to the relation

$$\alpha \approx 10^{-5} \beta \frac{\omega^{1/2}}{d}, \quad (4)$$

in SI units. The numerical parameter β is about 2.4 for perfectly smooth walls, but may become much larger than this for the rough walls characteristic of a naturally produced didjeridu.

If, as is typical of genuine instruments, the bore has nearly the shape of a flaring truncated cone with diameter d_1 at the blowing end and diameter d_2 at the open end, then the impedance coefficients have the more complicated forms^{26,27}

$$Z_{11}^D = -\frac{4j\rho c}{\pi d_1^2} \left[\frac{\sin(kL + \theta_2) \sin \theta_1}{\sin(kL + \theta_2 - \theta_1)} \right]$$

$$Z_{22}^D = +\frac{4j\rho c}{\pi d_2^2} \left[\frac{\sin(kL - \theta_1) \sin \theta_2}{\sin(kL + \theta_2 - \theta_1)} \right] \quad (5)$$

$$Z_{12}^D = Z_{21}^D = -\frac{4j\rho c}{\pi d_1 d_2} \left[\frac{\sin \theta_1 \sin \theta_2}{\sin(kL + \theta_2 - \theta_1)} \right],$$

where $\theta_1 = \arctan kx_1$ and $\theta_2 = \arctan kx_2$, with x_1 and x_2 being the axial distances of the ends 1 and 2 from the apex of the cone. Related expressions can also be derived for the input impedance Z_D when the bore is terminated by a given impedance,²⁷ in the present case the radiation impedance Z_R . These coefficients Z_{ij}^D must of course include the real parts due to viscous and thermal losses at the tube walls, as detailed in (4). If d_2 is the diameter and A_2 the area of the open end, then the radiation impedance is also a complex quantity which has the approximate form¹⁷

$$Z_R \approx 7 \times 10^{-4} \omega^2 + 0.5j \frac{\omega}{d_2} \quad \text{if } kd_2 \ll 1$$

$$\approx \frac{\rho c}{A_2} \approx \frac{500}{d_2^2} \quad \text{if } kd_2 \gg 1 \quad (6)$$

in SI units.

III. LIP-VALVE OPERATION

The lip valve is in reality a complex mechanism with many degrees of freedom. The player's lips are somewhat soft and flexible, their geometry is complicated, and they usually close completely for part of each vibration cycle, thus introducing very significant nonlinearity. As has been found with vocal-fold models, construction of any detailed model of the vibrating lips is a very major undertaking, and the appropriate model may well differ greatly from one player to another, or indeed during different types of performance by a single player. For these reasons, among others, it is preferable to develop a model that is as simple as possible, while still incorporating the major features of a real lip valve.

Experimental studies of players of brass-instruments players²⁸⁻³¹ have shown that the predominant lip motion at low frequencies can be described as that of “outward-swinging doors,” or the configuration (+, -) in the notation of Fletcher,³² where the first symbol indicates the action on the valve of a steady pressure on the upstream side and the second symbol the corresponding action of a downstream pressure. For fundamentals of high frequency, these studies of trumpet and trombone players show that a better approximation is that of “sliding doors” with configuration (+, +). Other studies have described the motion in terms of a combination of both types of motion²⁹ and some as surface waves propagating on the lips. To maintain generality we shall initially make no assumption about which of these configurations is a more appropriate approximation for the didjeridu and write the configuration symbol as (σ_1, σ_2) where, in the present case, $\sigma_1 = +1$ and σ_2 may be either +1 or -1 depending upon the lip model adopted.

Our own experimental studies of didjeridu players, described in detail in the companion paper,¹³ show that the lips are open for only about half of an oscillation cycle, that the lip opening is typically nearly elliptical, and that both elliptical diameters expand and contract to some extent together. The area S of the lip opening thus has approximately the form

$$S(t) = A \langle \cos \omega t - \cos \zeta \rangle^\mu, \quad (7)$$

where A is a constant, $1 \leq \mu \leq 2$, and the notation $\langle \dots \rangle$ is intended to imply that the enclosed expression is set to zero if it is negative. Here, ζ/π is the fraction of the period for which the lip valve is open, and since this is typically about half of the period for didjeridu players, $\zeta \approx \pi/2$, though this may vary somewhat with playing style. The expression (7) is usually referred to as a “transparent closure” model of the lip vibration, for obvious reasons.

While the opening and closing of the lips may not be exactly symmetrical in practice because of hysteresis effects in lip contact, we neglect this complication and write

$$S(t) = \sum_0^{\infty} s_n \cos n\omega t. \quad (8)$$

Consideration of (7) for the transparent closure model indicates that we should expect that $s_n \propto n^{-1}$, but this needs to be checked by experiment. Stroboscopic measurements on the lips of a didjeridu player, as shown in Fig. 8 of our experiment paper,¹³ allow the opening area to be plotted as a function of time, and from this the coefficients s_n can be determined. To a reasonable approximation, as shown in Fig. 9 of that paper, it is found that $s_n = n^{-2}s_0$, which differs from the predictions of the transparent closure model. The difference can be ascribed to a “soft” closure of the lips instead of the “sharp” closure when they pass through each other in this unrealistic model. The experimental result will be used in the calculations to follow, though it is possible that the player may be able to modify this relation to some extent by changing the tension in the lip muscles. Indeed, Fig. 9(c) of the experiment paper shows just such an effect in the “low-tongue” position, with the amplitude of higher harmonics s_n for $n > 5$ falling in amplitude even more rapidly than n^{-2} .

Since it is not the purpose of this paper to calculate lip motion in any detail, an outline based upon a transparent closure model referred to above in which, instead of a realistic nonelastic collision between the surfaces of the lips, it is assumed that each lip executes a sinusoidal motion and effectively passes through the other lip.¹¹ While this is clearly unrealistic, the properties of the collision can be adequately modeled by ascribing an appropriate damping to the motion of each lip. The outline of this model given below expresses it in one-dimensional form for simplicity.

Suppose that the upstream pressure is p_M and the downstream pressure p_D , as shown in Fig. 2(b), and that both act upon an effective lip area A . Let the lips be identical and have combined effective mass m and damping γ . If z measures the linear lip opening, then an adequate approximation to the lip behavior on this model is

$$m \frac{d^2 z}{dt^2} + \gamma \frac{dz}{dt} + m\omega_0^2(z - z_0) = A(\sigma_1 p_M + \sigma_2 p_D), \quad (9)$$

where ω_0 is the natural frequency of the lip vibration as determined by their mass and tension. In the swinging-door case, the parameters m and γ should be defined in terms of moments of inertia and centroid of pressure, but this makes no basic difference.

Discussion of a simplified version of this problem, in which the equilibrium opening is great enough that the valve does not close and the aperture width W is constant rather than varying in synchronism with the aperture height, has been given elsewhere.^{32,33} The conclusion is that the drone frequency must be a little above the didjeridu resonance in the case of a (+, -) lip valve and a little below the resonance in the case of a (+, +) valve. The player is presumably able to adjust the lip motion between these two configurations, as can trumpet players, though the total range of variation in the case of the didjeridu is only about a semitone

(6%). Informal observations of competent didjeridu players show that they typically raise the drone frequency when wishing to emphasize formants.

A general approach that properly includes all nonlinearities and treats finite vibration amplitudes is based upon the “method of slowly varying parameters” described by Bogoliubov and Mitropolsky.³⁴ In this approach it is assumed that the lip oscillation can be described by the expression

$$z = z_0 + a \sin(\omega_0 t + \phi), \quad (10)$$

and that both the amplitude a and the phase ϕ are slowly varying functions of time in the sense that they change by only a small amount over one period $2\pi/\omega_0$ of the oscillation. If (9) is rewritten in the form

$$\frac{d^2 z}{dt^2} + \omega_0^2(z - z_0) = g\left(z, \frac{dz}{dt}, p_M, p_D\right), \quad (11)$$

then it can be shown^{34,35} that

$$\frac{da}{dt} \approx \frac{1}{\omega_0} [[g \cos(\omega_0 t + \phi)]] \quad (12)$$

$$\frac{d\phi}{dt} \approx -\frac{1}{a\omega_0} [[g \sin(\omega_0 t + \phi)]], \quad (13)$$

where the notation $[[\dots]]$ is to be taken to imply that only terms varying slowly in time relative to $\sin \omega_0 t$ are to be retained. When the system settles into a steady state, the vibration amplitude a is constant and so too is $d\phi/dt$, implying that the new steady vibration frequency is $\omega_0 + d\phi/dt$. The qualitative results are the same as outlined above for the simple linearized model.

While detailed consideration of lip vibration is clearly necessary in a complete model for didjeridu sound production, creation of such a model is not the purpose of the present paper. Indeed, it will be sufficient to note that a model for lip vibration is possible and to specify its predictions by assuming reasonable values for the vibration amplitude and frequency. The frequency, as shown by (13), depends upon lip tension through the value of ω_0 , upon the didjeridu and vocal-tract impedances, and upon the nature of the lip motion as encoded in the parameters σ_2 and μ .

IV. SYSTEM ANALYSIS

In this section two approaches will be explored. The first, which is potentially able to predict the operation of the whole system in detail, relies upon the “harmonic-balance” approach to the behavior of the whole system. Unfortunately this approach is quite complex and necessarily numerical, so that the general trend of its conclusions is not readily apparent. The second approach reduces the whole system to the simplest possible approximation by considering only first-order terms. While the conclusions derived are certainly not quantitatively accurate, this approach does show clearly the general trends and, in particular, the way in which the spectrum of the output sound depends upon the vocal-tract parameters, which is the main purpose of the whole analysis.

In both cases the system to be modeled is that shown in Fig. 2. Part (a) of this figure gives a model for the player’s

lungs and vocal tract, while part (b) compresses this into a simpler system and includes the didjeridu as well. Finally, part (c) compresses the didjeridu subsystem further to give the electrical analog network that will be analyzed. The essential feature of this progressive simplification is that both the vocal tract and the didjeridu itself are ultimately represented by simple input impedances Z_V and Z_D , respectively, on the two sides of the lip valve. The values of these two quantities can be readily calculated, as has already been discussed.

A. Harmonic-balance approach

In the harmonic-balance approach³⁶ the procedure is as follows, after the relevant values of $Z_V(\omega)$ and $Z_D(\omega)$ have been evaluated. (i) Assume reasonable values for mouth pressure p_M , lip vibration frequency ω_0 , and lip vibration amplitude a . Take the pressure p_D at the input to the didjeridu tube to be zero. (ii) Calculate the volume flow through the lips as a function of time, using the Bernoulli equation. (iii) Take the Fourier transform of this flow and use it to calculate the Fourier components of the pressures p_M and p_D in the mouth and didjeridu input. (iv) Convert these pressures back into the time domain, or perhaps simply deal with the component at the fundamental frequency, which nearly matches the vibration frequency of the lips. (v) Use this driving pressure difference to recalculate the frequency and amplitude of the lip vibration, using either the “transparent closure” model, or preferably a more realistic model, for lip motion and the solution techniques detailed in Sec. III. (vi) Iterate steps (ii) through (v) until convergence is achieved. (vii) Calculate the resulting sound power radiation from the open end of the didjeridu tube.

It is clear that this is a rather complex though straightforward computational process, so that no simple general conclusions can be drawn except by examining behavior in many particular cases. For this reason such a detailed computation will not be carried out or discussed further at this time.

B. First-order approximation

Referring to Fig. 2(c), and denoting the lung pressure by p_L , the mouth pressure by p_M , and the input pressure to the didjeridu by p_D , the following equations can be written:

$$p_L - p_M = Z_V U, \quad (14)$$

$$p_D = Z_D U, \quad (15)$$

$$U = \left[\frac{2(p_M - p_D)}{\rho} \right]^{1/2} S(t), \quad (16)$$

where Z_V is the input impedance of the vocal tract at a position just inside the lips and Z_D is the input impedance of the didjeridu just outside the lips, $S(t)$ is the opening area of the lips, and U is the acoustic volume flow through the lips. Substituting (14) and (15) in (16) then gives

$$\begin{aligned} U &= \left(\frac{2}{\rho} \right)^{1/2} [p_L - (Z_V + Z_D)U]^{1/2} S(t) \\ &= \left(\frac{2}{\rho} \right)^{1/2} [p_L - ZU]^{1/2} S(t), \end{aligned} \quad (17)$$

where we have written $Z \equiv Z_V + Z_D$ for convenience. Suppose now that

$$U = \sum_0^{\infty} (u_n \cos n\omega t + v_n \sin n\omega t). \quad (18)$$

Recalling the Fourier expansion (8) for $S(t)$ and writing p_0 for the steady component of the mouth pressure p_M as given by (14),

$$\begin{aligned} &\sum_0^{\infty} (u_n \cos n\omega t + v_n \sin n\omega t) \\ &= \left(\frac{2}{\rho} \right)^{1/2} \left[p_0 - \sum_{n=1}^{\infty} Z_n (u_n \cos n\omega t + v_n \sin n\omega t) \right]^{1/2} \\ &\quad \times \sum_{m=0}^{\infty} s_m \cos m\omega t, \end{aligned} \quad (19)$$

where Z_n is the value of $Z_V + Z_D$ at frequency $n\omega$. Note that the introduction of p_0 has removed the zeroth-order term from the summation inside the parentheses because the didjeridu tube impedance $Z_D = 0$ when $\omega = 0$. In (19), since $Z_n = R_n + jX_n$ is a complex quantity, the symbol j must be interpreted as an operator meaning “advance the phase by $\pi/2$.”

As a reasonable approximation in the case of the didjeridu, we might assume that $p_0 \gg Z_n u_n, Z_n v_n$ for all n , so that the factor $[p_0 - \sum]^{1/2}$ on the right in (19) can be simply approximated by $p_0^{1/2} [1 - (1/2p_0)\Sigma]$, where in each case Σ represents the summation in this factor in (19). Since typically $s_1 \approx s_0 \gg s_n$ for all $n > 1$, we can also treat s_0, s_1, u_0, u_1 , and v_1 as being all of zeroth order, while $v_0 = 0$ and higher terms are all of first order or less. Even with this simplification, however, and limiting consideration to harmonics up to N , the result is an algebraically complicated, though sparse, $N \times N$ matrix equation containing nonzero elements M_{ij} for $j = i$ and $j = i \pm 1$.

While this matrix could be solved for a specific case, for a semiquantitative analysis it is appropriate to simplify (19) still further by assuming that $s_n \ll s_0$ for all $n > 0$, rather than just for $n > 1$. This carries with it the additional result that $u_n, v_n \ll u_0$ for all n . Physically this amounts to the assumption that, instead of closing in each cycle, the lips remain substantially open but vibrate in a nonlinear manner so as to generate all the terms s_n with magnitudes approximately proportional to $1/n^2$. Selection of terms in $\cos n\omega t$ and $\sin n\omega t$ in Eq. (19) then leads to the simpler results

$$\begin{aligned} u_n [(2p_0\rho)^{1/2} + s_0 R_n] + s_0 X_n v_n &= 2p_0 s_n \\ v_n [(2p_0\rho)^{1/2} + s_0 R_n] &= s_0 X_n u_n, \end{aligned} \quad (20)$$

which can be combined to give

$$u_n = \frac{2p_0 s_n [(2p_0 \rho)^{1/2} + s_0 R_n]}{[(2p_0 \rho)^{1/2} + s_0 R_n]^2 + s_0^2 X_n^2} \quad (21)$$

$$v_n = \frac{2p_0 s_0 s_n X_n}{[(2p_0 \rho)^{1/2} + s_0 R_n]^2 + s_0^2 X_n^2} \quad (22)$$

These two equations can be further combined to give the flow magnitude at frequency $n\omega$ as

$$(u_n^2 + v_n^2)^{1/2} = \frac{2p_0 s_n}{[(2p_0 \rho)^{1/2} + s_0 R_n]^2 + s_0^2 X_n^2} \quad (23)$$

which can be simply evaluated explicitly.

Since the quantity of interest is not the acoustic flow into the didjeridu from the lips, but rather the acoustic radiation from the open end of the tube, it is necessary to calculate the transfer function involved. Referring to the network analog shown in Fig. 2(b) and using a new notation with $U_1(n) \equiv u_n + jv_n$ and $U_2(n)$ being the acoustic flow from the tube into the radiation impedance Z_R , we can write

$$p_1(n) = Z_{11}^D U_1(n) - Z_{12}^D U_2(n) \quad (24)$$

$$p_2(n) = Z_{21}^D U_1(n) - Z_{22}^D U_2(n). \quad (25)$$

The minus signs in these two equations arise from the fact that both currents U_1 and U_2 have been taken to flow clockwise, as in Fig. 2(b), rather than symmetrically with respect to Z_{ij} as is sometimes the convention. At the open end of the tube $p_R = Z_R U_2$, so that (25) shows that

$$U_2(n) \approx \frac{Z_{21}^D}{Z_{22}^D + Z_R} U_1(n), \quad (26)$$

and this is the transfer function that is sought. Since normally $Z_R \ll Z_{22}^D$, with most of the resistive losses due to wall effects in the tube and the reactive term easily included as a simple end correction, this transfer function is approximately equal to Z_{21}^D / Z_{22}^D , which has maxima at the same frequencies as does the input impedance at the lip end of the tube.

The radiated acoustic power at frequency $n\omega$ is therefore

$$\Pi(n) = R_R(n\omega) U_2(n)^2 = R_R(n\omega) \left[\frac{Z_{21}^D}{Z_{22}^D + Z_R} U_1(n) \right]^2, \quad (27)$$

where R_R is the resistive component of the radiation impedance Z_R as given by (6) and $U_1(n) \equiv u_n$ in the previous analysis.

V. PREDICTIONS OF THE MODEL

The implications of this analysis can now be examined. The simplest case is that in which the didjeridu is a narrow uniform cylinder, for the impedance coefficients Z_{ij}^D for the tube are then easily expressed, as in (3), and the radiation resistance is simply proportional to the square of the frequency, as given by the lower range of (6). The expression for the radiated sound power, as given by (27), can then be evaluated explicitly once the relation of the drone frequency to the first resonance of the tube has been specified.

TABLE I. Model parameter values (high tongue).

Didjeridu tube length	L	1.3 m
Didjeridu tube diameter	d_1	40 mm
Loss coeff. for tube	β	12
First VT resonance	f_1	1.5 kHz
Second VT resonance	f_2	2.5 kHz
Q of VT resonances	Q	10
Charact. impedance	$(L/C)^{1/2}$	4 MPa s m ⁻³
Glottal open area		20 mm ²
Effective glottal length		5 mm
Blowing pressure in mouth	p_0	1 kPa
Max. area of lip opening	s_0	50 mm ²
Frequency rel. first resonance		1.03

As a first approximation, the small-signal result (27) can be used to predict the behavior of formants for such a cylindrical didjeridu. The result will not be accurate at low frequencies and applies to the high-tongue case in which the upper formants are emphasized at the expense of the drone fundamental. Because the vocal tract is a complex structure, this will be approximated by assuming that it possesses two resonant impedance minima above 1000 Hz, in this case at 1500 and 2500 Hz, respectively, these being separated by an impedance maximum. The tract can therefore be simply modeled as two series-resonant circuits in parallel, as shown in Fig. 1(d), with characteristic impedances appropriate to the cross section of the constricted upper part of the vocal tract. The third parallel impedance represents that of the glottis and lungs. The playing frequency was taken to be half a semitone (3%) above the fundamental resonance of the tube, this being fairly characteristic of actual playing in the case of emphasized formants. The other relevant parameters are given in Table I.

The results of this calculation are shown in Fig. 3 and agree very well with high-tongue spectra measured on

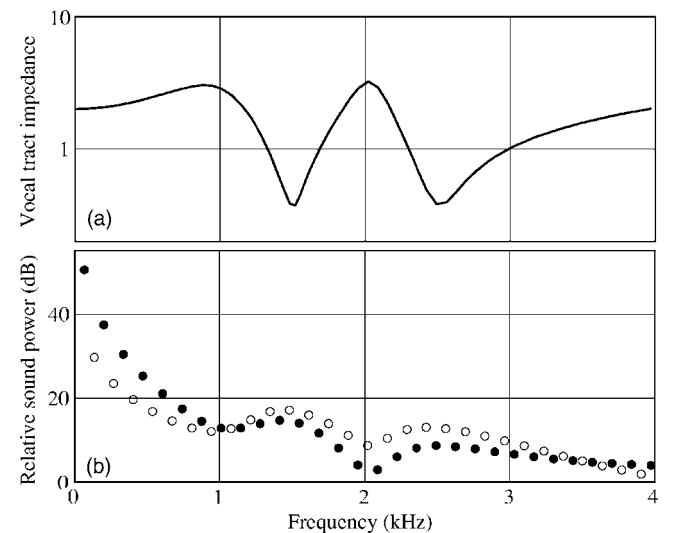


FIG. 3. (a) Vocal-tract impedance at the lips in units of MPa s m⁻³, with two simple resonances at 1.5 and 2.5 kHz, respectively. (b) Computed radiated spectrum for the case of a cylindrical didjeridu tube with other parameters as in Table I. Odd harmonics of the drone frequency are shown with solid circles and even harmonics with open circles.

skilled players, as shown in Fig. 6 of our experimental paper,¹³ which also used a cylindrical-pipe didjeridu. It is clear that there are two pronounced formants in the radiated sound spectrum, and that their center frequencies coincide with those of the two minima in the vocal-tract impedance. There is also a pronounced minimum at the frequency of the maximum in the tract impedance, as predicted. There are other interesting features of the spectrum illustrated in Fig. 3(b). Odd harmonics of the drone frequency are shown as filled circles and even harmonics as open circles. In the low-frequency range, odd harmonics are emphasized because of the maxima in the transfer function Z_{22}^D/Z_{12}^D close to those frequencies, but the fact that the drone frequency is a little above the first resonance means that the odd harmonics move progressively away from the transfer maxima, while the even harmonics move to closer matching and are consequently emphasized. An actual transfer of emphasis takes place at about 1000 Hz and another one in the reverse direction at about 3500 Hz. This effect is also clearly seen in experimental measurements on a cylindrical didjeridu.¹³ In particular cases where there is very near coincidence between harmonic frequencies and tube resonances in a particular spectral region, this effect may even lead to the production of an emphasized formant band of completely different origin.

In the calculation shown in Fig. 3, the playing frequency was chosen to be 1.03 times the first tube resonance, which is typical for high-tongue playing. There is in this case no exact coincidence between an upper harmonic and a tube resonance and therefore no single enhanced harmonic. A very skilled player can, however, adjust the playing frequency to produce such a coincidence and therefore a greatly enhanced amplitude for a particular harmonic, much as is done in the “throat singing” of Mongolia,³⁷ though this technique is not used in traditional playing.

A rather similar simplified analysis can be applied to the case of a didjeridu with a flaring bore, though the algebra is necessarily a little more complicated. The results are generally similar to those demonstrated above. Because, however, the frequencies of the low-frequency tube modes do not have a simple harmonic relationship, there are no clear frequency regions in which even or odd harmonics dominates. In addition, the low harmonics above the fundamental are generally weaker because their frequencies do not approximate those of tube resonances.

VI. VOCALIZED SOUNDS

For some of the sounds traditionally made on the didjeridu, the player vibrates his vocal folds as in singing, as well as maintaining the lip vibration. This technique is traditionally used to represent the cries of animals in a performance that is illustrating a particular story or describing a place. It is important to examine the nature of the resultant sound.

In the simplest case, the performer plays the drone at frequency ω_D and sings a note of rather higher frequency ω_V . Each of the vibrating valves modulates the air flow at its operating frequency, so that it has a form like

$$U = A \langle \sin \omega_D t - x_1 \rangle \langle \sin \omega_V t - x_2 \rangle, \quad (28)$$

where x_1 and x_2 determine the closing fraction for each valve and the notation $\langle \dots \rangle$ is meant to imply, as before, that only positive values of the quantity concerned are retained. The result is rather complicated and depends upon relative phasing of the two vibrations, but it is clear that the resulting flow will contain terms of essentially all frequencies $m\omega_D \pm n\omega_V$ where m and n are integers. The dominant terms will be those in which the values of m and n are small.

A particular example is the case in which the sung note is a musical fifth (frequency ratio 3:2) or more usually a musical tenth (frequency ratio 5:2) above the drone fundamental. The difference combination with $m=n=1$ in the first case or $m=2, n=1$ in the second then generates an impressive subharmonic drone of frequency $\omega_D/2$, together with all its harmonics.¹³ Most of the other sounds used in performance are complex and transient, so that little is to be gained by analyzing them in detail.

VII. DIDJERIDU QUALITY

Although the geometrical form of native didjeridus varies widely, and good players can produce striking performances on a simple length of cylindrical PVC pipe, these players are able to rate didjeridus in terms of quality. There do not appear to have been extensive studies of these ratings, but beginning attempts have been made.⁵⁻⁷ Amir,⁶ in particular, has produced a quality criterion based upon the levels of harmonics 2 to 5 relative to that of the fundamental, a didjeridu that is rated excellent by players having low levels for these harmonics.

Examination of the didjeridu preferences of different clans in Arnhem Land shows, however, that there are striking differences in the preferred didjeridu shape.³⁸ Clans from Eastern Arnhem land, including the Yolngu, adopt a playing style in which the use of rhythmic articulation and higher tube modes for accents is pronounced, and their preference is for almost cylindrical didjeridus. Clans from Western Arnhem Land, in contrast, make little use of these higher modes but exploit the timbre changes possible through the use of vocal-tract resonances, and their preference is for didjeridus with a flaring bore profile. Nontraditional didjeridu performers use a combination of both styles, since they give didjeridu sound a quality that is not heard in any other musical instrument.

It is easy to see the reasons for these preferences. A narrow tube with a nearly cylindrical bore will have strong lower resonances in nearly harmonic relationship and it will be relatively easy to produce the higher modes. The drone, too, will be strong in quality because the well-aligned resonances will reinforce the first few odd harmonics, a point that is important since the sensitivity of human hearing is reduced at the low frequency, 60–70 Hz, of the drone fundamental. The strong quality of the drone fundamental and its lower harmonics then makes less obvious the subtle timbre changes produced by formant variations in the higher harmonics. These physical consequences appear to explain the preferences of those playing in the Eastern Arnhem Land style.

The reverse is true for the Western Arnhem Land style and preferences. In a widely flaring tube, overtones other than the first are rather difficult to produce because of the rapidly decreasing quality factor of the tube resonances, while the misalignment of these lower resonances means that they do not reinforce the lower harmonics of the drone, making it much less prominent than in the Eastern style. This then has the consequence that formant bands produced by vocal-tract resonances, which do not rely upon reinforcement by the resonances of the instrument bore, are prominent in the total sound as desired.

VIII. CONCLUSIONS

The primary purpose of the present paper was to examine the relation between maxima and minima in the vocal-tract impedance and the formant bands observed as a prominent characteristic of expert didjeridu performance. While the model still requires development if it is to give an accurate quantitative description of the radiated spectrum, both the theoretical development and the computed spectra demonstrate a close relation between the emphasized formant bands in the sound spectrum and impedance minima of the vocal tract, as measured at the lips. Direct evidence for this correlation is presented in the accompanying experiment paper.

ACKNOWLEDGMENT

This work was supported by a grant from the Australian Research Council Discovery Program.

- ¹N. H. Fletcher, "Australian Aboriginal musical instruments: The didjeridu, the bull-roarer and the gumleaf," Proc. Stockholm Music Acoustics Conf. 2003, SMAC-03, pp. 201–204.
- ²T. A. Jones, "The didjeridu," Studies in Music (University of Western Australia) **1**, 23–55 (1967).
- ³T. A. Jones, "The yiraki (didjeridu) in Northeastern Arnhem Land: Techniques and styles," in *The Australian Aboriginal Heritage*, edited by R. M. Berndt and E. S. Phillips (Australian Society for Education through the Arts, in association with Ure Smith, Sydney, 1973), pp. 269–274.
- ⁴K. Neuenfeld, Ed., *The Didjeridu: From Arnhem Land to Internet* (John Libbey, Sydney, 1997).
- ⁵N. Amir, "A study of the didjeridu: Normal modes and playing frequencies," Proc. Internat. Symp. Musical. Acoust. ISMA-01, Perugia, Italy, 2001, pp. 95–98.
- ⁶N. Amir, "Harmonics: What do they do in the didjeridu?," Proc. Stockholm Music Acoust. Conf., SMAC-03, Stockholm, Sweden 2003, pp. 189–192.
- ⁷R. Caussé, G. Baudouin, and B. Sluchin, "An investigation on 'tonal' and 'playability' qualities of eight didgeridoos, perceived by players," Proc. Internat. Symp. Musical. Acoust. ISMA-04, Nara, Japan 2004, pp. 261–264.
- ⁸N. H. Fletcher, "Acoustics of the Australian didjeridu," Australian Aboriginal Studies **1**, 28–37 (1983).
- ⁹N. H. Fletcher, "The didjeridu (didgeridoo)," Acoust. Aust. **24**, 11–15 (1996).
- ¹⁰G. C. Wiggins, "The physics of the didgeridoo," Phys. Bull. **39**, 266–267 (1988).
- ¹¹L. Hollenberg, "The didjeridu: Lip motion and low frequency harmonic generation," Aust. J. Phys. **53**, 835–850 (2000).
- ¹²N. Fletcher, L. Hollenberg, J. Smith, and J. Wolfe, "The didjeridu and the

- vocal tract," Proc. Internat. Symp. Musical. Acoust. ISMA-01, Perugia 2001, pp. 87–90.
- ¹³A. Z. Tarnopolsky, N. H. Fletcher, L. C. L. Hollenberg, B. D. Lange, J. Smith, and J. Wolfe, "Vocal tract resonances and the sound of the Australian didjeridu (yidaki) I. Experiment," J. Acoust. Soc. Am. **119**, 1194–1204 (2006).
- ¹⁴A. Tarnopolsky, N. Fletcher, L. Hollenberg, B. Lange, J. Smith, and J. Wolfe, "The vocal tract and the sound of a didgeridoo," Nature (London) **436**, 39 (2005).
- ¹⁵L. L. Beranek, *Acoustics* (Acoustical Society of America, New York, 1986), Chap. 3.
- ¹⁶N. H. Fletcher, *Acoustic Systems in Biology* (Oxford University Press, New York, 1992), Chaps. 8 and 10.
- ¹⁷N. H. Fletcher and T. D. Rossing, *The Physics of Musical Instruments*, 2nd ed. (Springer, New York, 1998), Secs. 8.2, 8.3, 8.15.
- ¹⁸E. R. Weibel, *Morphometry of the Human Lung* (Springer, Berlin, 1963).
- ¹⁹K. Ishizaka, M. Matsudaira, and T. Kaneko, "Input acoustic-impedance measurements of the subglottal system," J. Acoust. Soc. Am. **60**, 190–197 (1976).
- ²⁰A. C. Jackson, B. Suki, M. Ucar, and R. Habib, "Branching airway network models for analyzing high-frequency lung input impedance," J. Appl. Physiol. **75**, 217–227 (1993).
- ²¹S. Mukai, "Laryngeal movement during wind instrument play," J. Otolaryngol. Japan **92**, 260–270 (1989).
- ²²S. Mukai, "Laryngeal movement while playing wind instruments," Proc. Int. Sympos. Musical Acoust. ISMA-92, Tokyo, Acoust. Soc. Japan 1992, pp. 239–242.
- ²³J. Wolfe, J. Smith, G. Brielbeck, and F. Stocker, "A system for real time measurement of acoustic transfer functions," Acoust. Aust. **23**, 19–20 (1995).
- ²⁴J. Epps, J. R. Smith, and J. Wolfe, "A novel instrument to measure acoustic resonances of the vocal tract during phonation," Meas. Sci. Technol. **8**, 1112–1121 (1997).
- ²⁵J. Wolfe, A. Z. Tarnopolsky, N. H. Fletcher, L. C. L. Hollenberg, and J. Smith, "Some effects of the player's vocal tract and tongue on wind instrument sound," Proc. Stockholm Music Acoust. Conf., SMAC-03, Stockholm, Sweden, 2003, pp. 307–310.
- ²⁶N. H. Fletcher, *Acoustic Systems in Biology* (Oxford University Press, New York, 1992), Appendix B.
- ²⁷H. F. Olson, *Acoustical Engineering* (Van Nostrand, New York, 1957), pp. 100–123.
- ²⁸S. Yoshikawa, "Acoustical behavior of brass player's lips," J. Acoust. Soc. Am. **97**, 1929–1939 (1995).
- ²⁹S. Adachi and M. Sato, "Trumpet sound simulation using a two-dimensional lip vibration model," J. Acoust. Soc. Am. **99**, 1200–1209 (1996).
- ³⁰D. C. Copley and W. J. Strong, "A stroboscopic study of lip vibrations in a trombone," J. Acoust. Soc. Am. **99**, 1219–1226 (1996).
- ³¹N. H. Fletcher and T. D. Rossing, *The Physics of Musical Instruments*, 2nd ed. (Springer, New York, 1998), Sec. 13.3.
- ³²N. H. Fletcher, "Autonomous vibration of simple pressure-controlled valves in gas flows," J. Acoust. Soc. Am. **93**, 2172–2180 (1993).
- ³³A. Tarnopolsky, N. H. Fletcher, and J. C. S. Lai, "Oscillating reed valves—An experimental study," J. Acoust. Soc. Am. **108**, 400–406 (2000).
- ³⁴N. N. Bogoliubov and Y. A. Mitropolsky, *Asymptotic Methods in the Theory of Nonlinear Oscillations* (Gordon and Breach, New York, 1961).
- ³⁵N. H. Fletcher and T. D. Rossing, *The Physics of Musical Instruments*, 2nd ed. (Springer, New York, 1998), Chap. 5.
- ³⁶J. Gilbert, J. Kergomard, and E. Ngoya, "Calculation of the steady-state oscillations of a clarinet using the harmonic balance technique," J. Acoust. Soc. Am. **86**, 35–41 (1989).
- ³⁷S. Adachi, "An acoustical study of sound production in biphonic singing, Xöömij," J. Acoust. Soc. Am. **105**, 2920–2932 (1999).
- ³⁸S. Knopoff, "Accompanying the dreaming: Determinants of didjeridu style in traditional and popular Yolngu song," in *The Didjeridu: From Arnhem Land to Internet*, edited by K. Neuenfeld (John Libbey, Sydney, 1997), pp. 39–67.