ACOUSTICAL COUPLING BETWEEN LIP VALVES AND VOCAL FOLDS

Joe Wolfe and John Smith

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

J.Wolfe@unsw.edu.au

ABSTRACT: In normal speech and singing, the standing waves produced in the upper vocal tract are thought to have relatively little effect on vocal fold vibration. We demonstrate the effect of acoustic waves on vocal fold motion. The waves, whose magnitudes are comparable with those produced by vocalisation, are produced by playing a pipe in the manner of a didjeridu. We monitor vocal fold and lip motion by measuring their electrical admittance through the skin, and compare them with the radiated sound. In the presence of deliberate vocalisation, interesting heterodyne effects are produced, which are visible in all three signals. When the folds are relaxed, or in the configuration used for whispering, standing waves produced by playing the pipe produce vibrations in the vocal folds whose magnitudes are comparable with those associated with active vocalisation.

1. INTRODUCTION

In the widely used source-filter model for speech, the vocal tract is regarded as a filter whose frequency dependence modifies the spectrum produced by the vibration of the vocal folds [1, 2]. The frequency of the vibration of the folds (typically 100 to 300 Hz) usually lies below those of the resonances in the vocal tract (the first is typically 300 to 800 Hz). The frequency of vocal fold vibration is determined by mechanical parameters of the folds, the average subglottal pressure applied across them and aero-acoustic effects. Of course, the acoustic pressure waves in the vocal tract also exert force on the vocal folds, and thus may affect their motion [3,4] but this is usually thought to be a rather smaller effect. In brief, the vocal folds 'drive' acoustic waves in the vocal tract.

In lip-reed musical instruments, such as trumpet, tuba, etc, the lips of the player regulate air flow into the instrument which, like the vocal tract, has a number of resonances [5]. The lips of the player are therefore somewhat analogous to the vocal folds in the case of the voice. However, in such instruments, various instrument and lip parameters are usually chosen so that the standing waves in the bore of the instrument 'drive' the player's lips, which thus oscillate at a frequency close to one of the resonances of the bore. In brief, acoustic standing waves in the instrument bore 'drive' the lips.

Is it possible for pressure waves in the vocal tract to control the motion of the vocal folds in an analogous manner? In normal singing, this seems unlikely: most forms of singing require singers to control pitch and phonemes independently: in other words, they must control the frequency of the fold vibration and the tract resonances independently. The tract resonances have a relatively low Q factor and, for male voices especially, lie at frequencies well above that of the fold vibration.

Large amplitude pressure waves are produced in the vocal tract when playing musical wind instruments. However, vocalisation (the periodic opening of the vocal folds whilst playing) is not usually used in performance on most wind instruments. On the didjeridu, however, vocalisations are an important performance technique. So a study of vocalisation in didjeridu performance may therefore give information about the extent to which the vocal folds may be influenced by pressure waves.

Here we report briefly the results of a study in which vocal fold motion is partly or completely controlled by acoustic waves in the tract produced by a 'didjeridu' – in this case a plastic pipe being used in the manner of a didjeridu. The vocal fold and lip motion were studied simultaneously by measuring the electrical admittance between one pair of electrodes placed either side of the neck, at the level of the vocal folds, and that between another pair placed either side of the lips. The output sound and the sound pressure in the player's mouth were also measured, simultaneously.

One of the features of didjeridu performance is the production of heterodyne components when the player vocalises at a pitch different from that of the instrument [6]. Here we show the lip, vocal fold and sound signals involved in such heterodyne production, and show that the magnitude of vibrations of the folds due to standing waves in the vocal tract may be comparable with the magnitude of vibrations produced by the voice itself. We also show the effect of the waves in driving passive vocal folds.

2. MATERIALS AND METHODS

Two disc electrodes (33 mm diameter) of an electroglottograph (EGG) (model EG-2, Glottal enterprises, Syracuse, NY) were coated with conducting gel and positioned conventionally on either side of the throat, at the level of the vocal folds. They monitored the aperture of the glottis, the gap between the vocal folds. Another pair was positioned on either side of the lips, as shown in Fig. 1, to monitor the contact between the lips. All electrodes were held firmly in place with Velcro bands about the neck and head. The EGG supplied a current at 2 MHz and the output signals from each channel correspond to the admittance between the electrodes. Closing the glottis or closing the lip aperture in each case increases the admittance, so the trace rises as the contact between the folds or lips is increased. Is there significant electrical crosstalk between the two pairs of electrodes? When the mouth is open, the admittance signal across the lips produced by vocalisation is too small to measure. In contrast, the admittance signal across the relaxed vocal folds produced by lip vibration can be measured: it is typically 5% of that measured simultaneously at the lips. However, this vocal fold signal is not noticeably changed by disconnecting the electrodes from the lips: consequently the coupling is acoustical, not electrical. Nevertheless, the lip electrodes were disconnected when not required to eliminate the possibility of any crosstalk.

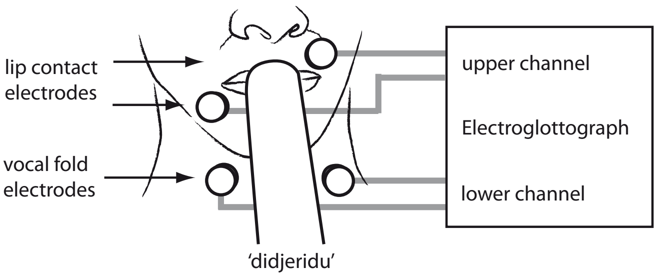

Figure 1. A schematic diagram showing the approximate positions of the electroglottograph (EGG) electrodes.

The aim of the experiment was to investigate the possible effects of heterodyne production and to observe the effects of large amplitude waves on the vocal folds. This only required a player who could vocalise reliably at the desired pitch whilst playing; consequently the experiment was conducted using one of the authors (JW). We were able to detect reliably the features of interest. An experienced player could presumably produce louder vocalisations and consequently stronger heterodyne components. However, the aim did not involve determining any parameters typical of didjeridu playing; and consequently there was little to be gained by measuring additional subjects.

A study involving didjeridu playing styles would normally use traditional instruments, however in this project we were only concerned with acoustical properties rather than musical significance. Furthermore, traditional didjeridus can carry significant spiritual significance that cannot be known by the investigators, and this can be problematical. Consequently, two simple cylindrical plastic pipes were used as substitutes. One had a length of 121 cm and an internal diameter of 34 mm – values typical of a didjeridu. The other had a length of 52 cm and an internal diameter of 26 mm. The musical quality of such pipes is not ranked highly by players [7]. However, this is not significant in this experiment. The sound was measured with an electret microphone positioned 10 cm from the bell, on the axis of the instrument. Sound pressures in the mouth were measured using a calibrated microphone (Bruel and Kjær Deltatron 1/4" type 4944A) with a Nexus conditioning amplifier.

3. RESULTS AND DISCUSSION

Heterodyne components

Figs. 2 to 4 show spectra of the radiated sound, the electrical admittance measured across the lips and the electrical admittance measured across the vocal folds, all measured simultaneously. In each case the subject vocalised at a consonant musical interval above the fundamental of the pipe, so as to produce simple heterodyne tones.

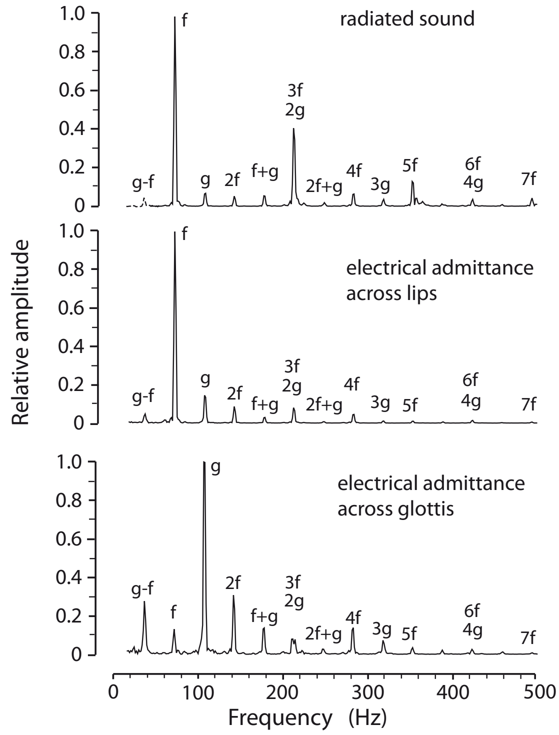

Figure 2. The spectra of the radiated sound, the electrical admittance across the lips and that across the glottis. The subject vocalises at a frequency g (≈106 Hz) that is 3/2 times (i.e. a perfect fifth above) the fundamental frequency of the long pipe f (≈71 Hz). The spectra were calculated from a series of 33072 samples lasting 750 ms using a Hann window. For clarity the sound spectrum only has been increased by a factor of 10 for frequencies below f (indicated by a dashed line).

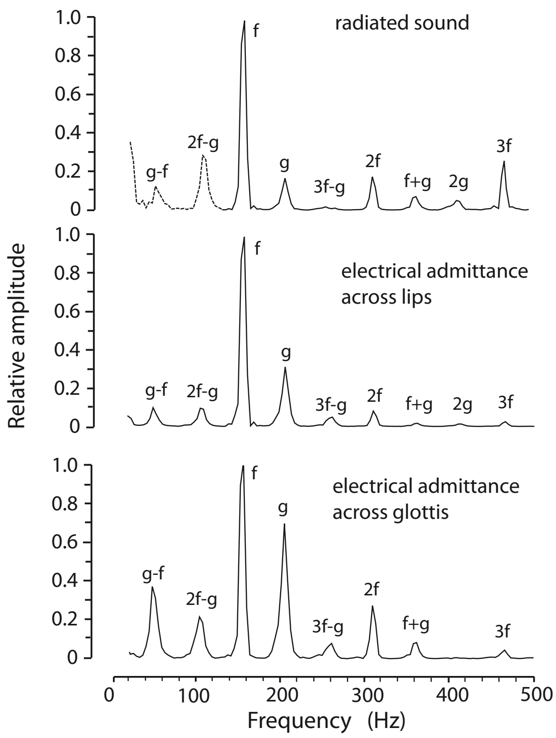

Figure 3. The spectra when the subject vocalises at a frequency g (≈206 Hz) that is 4/3 times (i.e. a perfect fourth above) the fundamental frequency of the short pipe f (≈155 Hz). The spectra were calculated from a series of 12000 samples lasting 272 ms using a Hann window. See caption for figure 2 for more details.

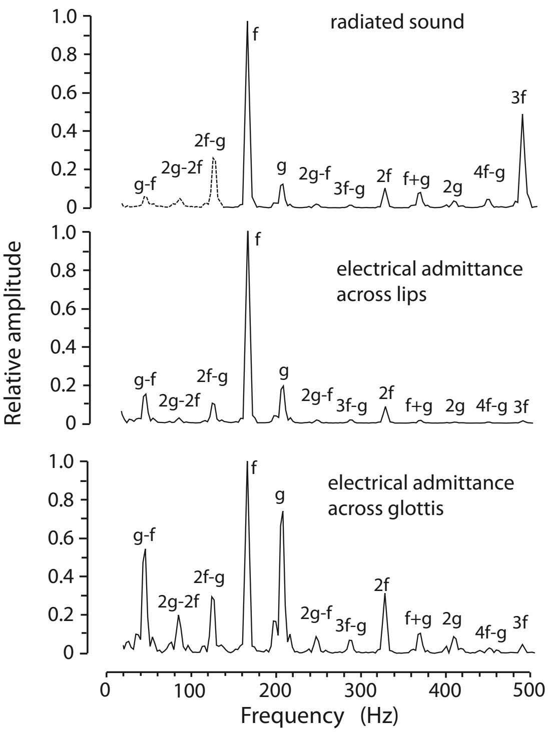

Figure 4. The spectra when the subject vocalises at a frequency g (≈205 Hz) that is 5/4 times (i.e. a major third above) the fundamental frequency of the short pipe f (≈164 Hz). The spectra were calculated from a series of 15870 samples lasting 360 ms using a Hann window. See caption for figure 2 for more details.

In Fig. 2, the longer pipe is used and the subject 'sings' a note (with fundamental frequency *g*) a musical fifth above that of the instrument (fundamental frequency f): *i.e.* $g = 3f/2$. In this case, the difference frequency $g - f = f/2$, and consequently occurs an octave below *f*. In Fig. 3, the shorter pipe is used and the subject vocalises at a perfect fourth so $g = 4f/3$, giving a difference frequency $g - f = f/3$, corresponding to a frequency one octave plus a fifth below the fundamental of the instrument. In Fig. 4, the subject vocalises at a major third so $g = 5f/4$, giving a difference frequency $g - f = f/4$, corresponding to two octaves below the fundamental of the instrument.

In these three cases, the air flow into the instrument is modulated by the periodic, but non-sinusoidal, motion of both the lips and vocal folds. The open areas of the lips (L) and glottis (G) can be modelled as $S_L = \Sigma a(n) \sin(2\pi nft)$ and $S_G = \Sigma b(m) \sin(2\pi mgt)$ respectively, where $a(n)$ and $b(m)$ are the amplitudes of the Fourier components, *n* and *m* integers. In a very simple model that neglects the effects of the vocal tract impedance, the flow through the lips is proportional to the product $S_G S_L$ and so has components at all frequencies $nf \pm mg$ [6,8,9]. For the simple, musically consonant cases shown here, these terms are all harmonics of the difference frequency $g - f$, and virtually all these are present in all the spectra shown in Figs. 2 to 4. The difference frequency $g - f$ itself is rather weak in the sound signal, in part because lower frequencies are less well radiated from a pipe than are higher frequencies.

In the spectra, there are some similarities between the lip signal and the radiated sound. Although the fundamental frequency of the lips is largely determined by the lowest resonance of the pipe, the motion of the lips determines the flow of air into the pipe and thus strongly influences the sound that is produced. The oscillating air flow through the lips produces sound waves, with comparable amplitude, that travel in both directions: into both the pipe and the vocal tract. While the spectrum of the sound inside the mouth was not measured in this set of measurements, one would expect its spectrum also to share features with that of the lip motion, as modified by the resonances in the vocal tract itself [8,10]. However, there is one systematic difference between lip and sound spectra. The pipe is nearly closed by the lips, so the resonances of the pipe fall at frequencies that are close to odd harmonics of the fundamental. Consequently, at the frequencies shown in these figures, the odd harmonics of the lip motion lie close to resonances of the pipe, and so are well matched to the radiation field. Hence the first few odd harmonics in the output sound (f, $3f$, $5f$) are stronger than the even harmonics ($2f$, $4f$).

In Fig. 2, the vocal fold signal shows, as expected, a strong component at the vocalisation frequency $g = 3f/2$. In this example, the ratios of the amplitudes of components at frequencies *f* and *g* are inverted between the lip signal and the fold signal. The folds thus influence the fundamental of the lip motion in approximately the same proportion that the lip motion influences the fold motion.

The experiments represented in Figs. 3 and 4 show frequency ratios $g/f = 4/3$ and $5/4$. A shorter pipe was used because the subject found it difficult to produce a powerful vocal signal at $5f/4$ for the longer pipe (about F#2). In part, this is because it is close to the lower end of his vocal range, and in part it was because of the difficulty of controlling the vocal folds in the presence of low frequency interference from the instrument sound in the tract. On the short pipe, $5f/4$ (A3) fell in a range in which he could sing loudly. Nevertheless, it is interesting to see that the influence of the lip signal (at frequency *f*) on the vocal signal (at frequency *g*) is much stronger than the reverse. For no combination of pipes, notes and blowing pressure was the subject able to produce a vocalisation signal whose influence on the lip signal was much stronger than the converse. This is perhaps not surprising: although the subject had the impression of singing loudly at frequency *g*, the lips are driven by a high Q resonator at *f*. The lips are strongly coupled to the mouth as well as to the pipe, and so produce a large amplitude pressure wave inside the mouth.

Figs 5 and 6 show (in the time domain) the sound pressure inside the mouth and the admittance at the vocal folds while the subject plays the shorter pipe. Lip electrodes were not used for these measurements to eliminate the possibility of crosstalk. For comfortable (neither loud nor soft) playing levels, the sound pressure level inside the mouth varied between about 130 and 145 dB with respect to 20 μPa and varied only slightly with position in the mouth, which is not surprising for large wavelengths. (The sound level in such experiments was also a few dB larger for the longer pipe than for the shorter.) How do these sound levels compare with those measured inside the mouth while singing? The subject was asked to sing and to produce a sound level 10 cm outside the mouth similar to that produced by the played pipe, measured 10 cm from its end. Singing "oo" (mouth nearly closed) produced a sound level in the mouth of 136 to 140 dB. Humming produced similar sound levels inside the mouth (around

140-145 dB), but the sound level measured outside the mouth and nose was 10 to 12 dB lower than for singing.

So figs. 2 to 4 show clear examples of how the motion of the vocal folds, even during deliberate vocalisation, can be strongly influenced by independently generated pressure waves in the tract. This observation has potential importance in understanding source-filter interactions in speech, singing and other contexts. What happens when the vocal folds are subject to large amplitude pressure waves when they are relaxed rather than vocalising?

Standing waves and 'passive' vocal folds

Fig. 5 shows how the vocal folds can be affected even when the player is not actively vocalising. In this typical example, the subject initially plays the pipe at a comfortable level while "whispering into the instrument" i.e. positioning his vocal folds in the position used for whispering. Although the subject is not actively vocalising, the amplitude of vibration recorded from the vocal folds is similar to that measured for active vocalisation. The vocal folds are then moved to a position that the subject reports as relaxed position, without consciously changing the other articulators. The vocal fold signal decreases, but does not become negligible, indicating the vocal fold contact area is still varying. Acoustically driven vibration of the vocal folds has been measured previously via laryngeal endoscopy [11].

Fig. 6 shows an example in which the subject vocalised deliberately, while playing a note on the pipe at comparable amplitude, so as to produce clear interference beats. All scales are the same as in Fig. 5, and there were no changes in the apparatus between the measurements. In the first part of the trace shown in Fig. 6, the two frequencies differ by a few Hz, producing the interference beats that appear as amplitude fluctuations in both the sound pressure and the vocal fold admittance. The subject then tunes the vocal fold vibration to match that of the lips, while maintaining a similar vocal effort. The beats seen in the first part of the trace suggest that the two signals have comparable size.

Now compare the first part of the traces in Fig. 5 with those

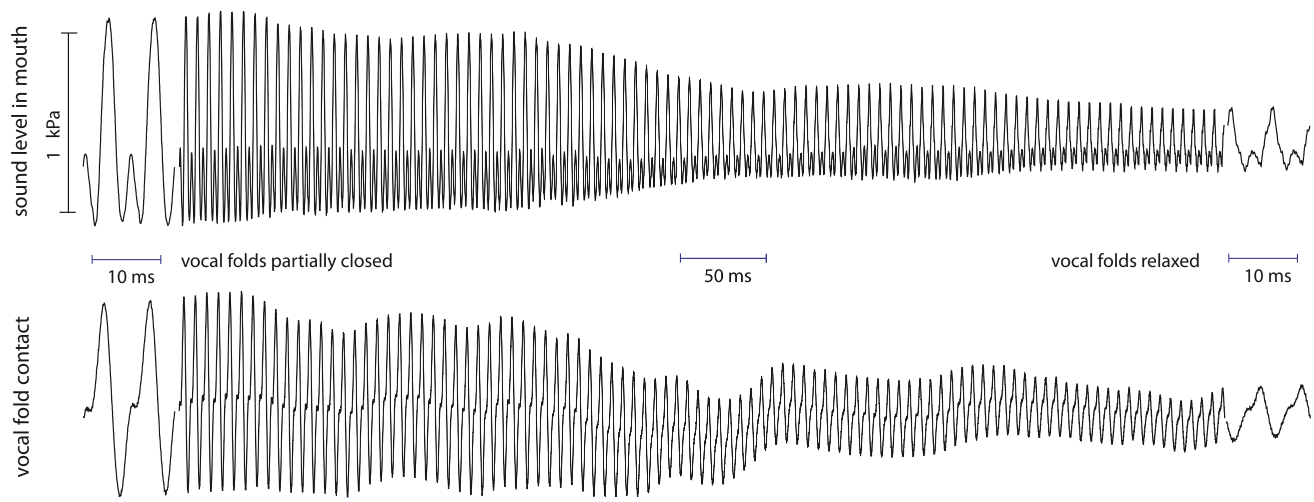

Figure 5. Oscillograms showing the sound pressure in the mouth (upper trace) and the admittance measured across the vocal folds (lower). The initial and final two cycles are shown on an expanded time scale. During this segment, the subject moves his vocal folds from the whispering configuration (vocal folds partially closed) to the relaxed configuration.

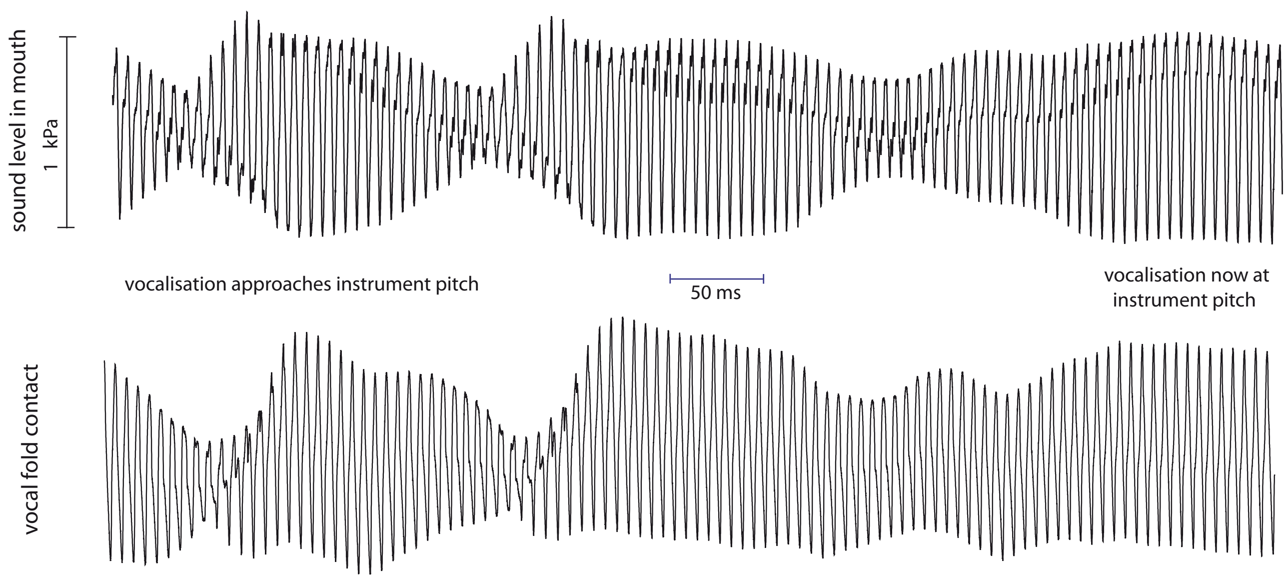

Figure 6. Oscillograms showing the sound pressure in the mouth (upper trace) and the admittance measured across the vocal folds (lower). During this segment, the subject adjusts the pitch of his vocalisation to match that of the instrument. The scales for vocal fold contact and pressure are identical to those of Fig. 5.

in the first part of Fig. 6. In both these cases, the acoustic waves in the mouth have similar amplitude. The variation in vocal fold contact also has similar magnitude: in the presence of these acoustical waves, simply putting the vocal folds in the whispering position produces a vocal fold signal comparable in size with that produced by deliberate vocalisation. The vocal fold signal effectively measures contact area and it is difficult to relate this quantitatively to glottis size. Nevertheless, this observation suggests that the vibration of the vocal folds when 'driven' by an acoustic signal in the whispering position can be comparable in magnitude with those produced by deliberate vocalisation.

In simple models, Rothenberg [3] and Titze [4] have considered the acoustic load of the vocal tract on the motion of the folds. The current study indicates that the effect of acoustic waves in the tract on vocal fold motion is considerable. One would expect this effect to be greatest when a harmonic of the fold motion lay near a resonance of the vocal tract. We have reported consistent examples of this vocal tract tuning in two classes of singing [12-14, see also 15]. The impedance of the tract 'seen' by the vocal folds has a large imaginary component (i.e. pressure and flow out of phase). This component changes sign as the frequency passes through the resonance, going from inertive below the resonance to compliant above. The observation that strong pressure waves can drive the vocal folds suggests a mechanism whereby singers learn the technique of vocal tract tuning: perhaps it is physically easier to sing when the resonance is close to the fold vibration frequency.

CONCLUSIONS

Electrical admittance measured across the lips is an effective way to monitor lip motion in lip-valve wind instruments. Measurements of vocalisations at harmonic intervals show the expected heterodyne effects, not only in the sound, but also in the lip motion and especially in the vocal fold motion. We demonstrate that pressure waves with amplitudes around one kPa in the mouth can drive 'passive' vocal folds at amplitudes comparable with those used in deliberate vocalisation.

ACKNOWLEDGMENT

We thank the Australian Research Council for their support.

REFERENCES

[1] G. Fant, *Acoustic Theory of Speech Production*, (Mouton and Co. Gravenhage 1960)

[2] J. Clark, C. Yallop and J. Fletcher, *An Introduction to Phonetics and Phonology,* (Basil Blackwell. Oxford 2007)

[3] M. Rothenberg, "An interactive model for the voice source" *Quarterly Prog. Status Report*, Dept Speech, Music and Hearing, KTH, Stockholm, **22**, 1-17 (1981).

[4] I. R. Titze, "The physics of small-amplitude oscillations of the vocal folds" *J. Acoust. Soc. Am.* **83**, 1536-1552 (1988)

[5] N.H. Fletcher and T.D. Rossing, "The physics of musical instruments" (Springer-Verlag, 1998)

[6] N. Fletcher, "The didjeridu (didgeridoo)" *Acoustics Aust*. **24**, 1-11 (1996)

[7] J. Smith, G. Rey, P. Dickens, N. H. Fletcher, L. C. L. Hollenberg and J. Wolfe, "Vocal tract resonances and the sound of the Australian didjeridu (yidaki): III. Determinants of playing quality" *J. Acoust. Soc. Am*. **121**, 547–558 (2007)

[8] A. Tarnopolsky, N. Fletcher, L. Hollenberg, B. 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)

[9] N. Fletcher, L. Hollenberg, J. Smith, A. Tarnopolsky and J. Wolfe, "Vocal tract resonances and the sound of the Australian didjeridu (yidaki) II: Theory", *J. Acoust. Soc. Am.* **119**, 1205-1213 (2006)

[10] A. Tarnopolsky, N. Fletcher, L. Hollenberg, B. Lange, J. Smith and J. Wolfe, "The vocal tract and the sound of the didgeridoo", *Nature* **436**, 39 (2005)

[11] S. Hertegard, H. Larsson and S. Grandqvist "Vocal fold resonances at high and low pitch tuning", *Proc. Stockholm Music Acoustics Conference (SMAC 03)*, (R. Bresin, ed) Stockholm, Sweden. 459–461 (2003)

[12] E. Joliveau, J. Smith and J. Wolfe, "Tuning of vocal tract resonances by sopranos" *Nature* **427**, 116 (2004)

[13] E. Joliveau, J. Smith and J. Wolfe, "Vocal Tract Resonances in Singing: The Soprano Voice", *J. Acoust. Soc. Am.* **116**, 2434-2439 (2004)

[14] N. Henrich, M. Kiek, J. Smith and J. Wolfe, "Resonance strategies in Bulgarian women's singing" *Logopedics Phoniatrics Vocology* **32**, 171-177 (2007)

[15] J. Sundberg, *The Science of the Singing Voice*, (Northern Illinois Univ. Press., De Kalb, IL 1987)