Estimating pressure and flow at the glottis in a vocal tract-like duct from microphone measurements at the mouth

Hugo Lehoux, Andre Almeida, Noel Hanna, Joe Wolfe and John Smith

School of Physics, UNSW Sydney

a.almeida@unsw.edu.au

Abstract

We describe a method to deduce pressure and flow at the glottis from acoustic measurements near or downstream from the lips, using phonation into a cylindrical duct with three microphones. In this paper this method is tested on hard one-dimensional ducts, also reporting the precision of using 3D printed models of vocal tract area functions. Combining all the approximations yields results that are poor for the frequency range needed for phonetics, but they perform better at the frequencies of the first few voice harmonics.

Index Terms: vocal folds, physics, acoustic propagation

1. Introduction

One of the long-term goals of voice research is to calculate the pressure and flow at the glottis from measurements made through the lips without prior assumptions about the glottal flow. Several approaches to determining acoustic variables at the glottis of a speaker or singer have been reported: Rothenburg measured flow at the mouth using a mask [1] and Alipour [2] measured flow velocity and pressure directly on excised larynges of mammals using for instance a hot-wire probe, while others calculate these quantities using inverse filtering that requires some assumptions about the shape of the glottal flow [3,4], or a priori estimations of the vocal tract transfer function [5]. Inverse filtering has the advantage of being a non-invasive technique, but the assumptions about the glottal flow may be sometimes too strong for studies of the 3-dimensional motion of real vocal folds.

This paper suggests an alternative technique based on a method of acoustic impedance measurement that uses the pressure information provided by two or more microphones along a cylindrical tube (an impedance head) to deduce pressure and flow anywhere within that tube. The same formulae that are applied in that acoustic measurement can, in principle, be applied to determine the pressure and flow anywhere along the duct being measured, provided that the duct geometry is known and the frequencies of interest and the load geometry are such that the acoustic propagation can be considered as one-dimensional (1d). This in turn allows the determination of a (frequency-dependent) transfer matrix relating the Fourier magnitudes of pressures and flows at any two positions in the duct.

To test this approach, acoustic propagation in a known 1d duct is applied to the calculation of the flow and pressure in artificial ducts. The artificial ducts (*e.g.* Fig 1) have an area function $A(x)$ (cross section A as a function of distance x from the glottis) calculated from scans of real vocal tracts for given vowel articulations. 3D printed vocal tract models have been used to model the acoustic characteristics of vocal tracts using excitation at the glottis and to test measurements of transfer functions (e.g. [6,7,8,9]). Sometimes the quality of printing is an uncontrolled variable. In the experiments described here, known flow and pressure can be input at the 'glottis' end using one impedance head. This can be compared with the calculations using the propagation model and measurements made with a second impedance head at the 'lip' end. As such we can report on the acoustic quality of 3D-printed vocal tracts, benchmarked against some simple geometries. (We shall use 3D for the fabrication technique, and 3d for three-dimensional in other senses.)

For most of the measurements, a loudspeaker and a synthetic broad-band signal is used for spectral measurements. In other measurements, the voice itself or a synthesised glottal flow is used.

2. Materials and methods

2.1. 3D printed one dimensional tracts

1d vocal tract geometries were 3D-printed on a Cubicon Single Plus (3DP-310F) printer in PLA filament (Filaform and Verbatim 1.75 mm) based on previously published area functions calculated from MRI scans of real vocal tracts [10]. The area functions were adjusted slightly to match the lip end smoothly to the 26.2 mm diameter of one impedance head, and to the 7.8 mm radius at the glottis to match the other impedance head. Most tracts were printed in two longitudinal halves (see Fig. 1), in order to inspect the accuracy of the print. The two halves were later bonded together by applying an epoxy resin to the join between the two external surfaces to prevent liquid flowing into the airway.

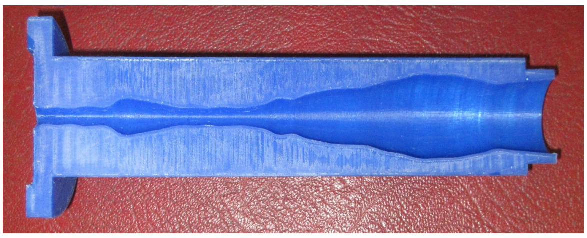

Figure 1: *Half of a 3D printed vocal tract model for the vowel /a/ from the data of Story et al. (1996). Note the left (glottis) side has been adjusted to fit the 7.8 mm diameter impedance head, and the right (lip) side has been adjusted to fit the 26.2 mm internal diameter impedance head.*

A few simple geometries such as cylinders and conical tract shapes were also printed and prepared in the same way and some were also printed in one piece for comparison.

2.2. Experimental setup

The experimental setup uses two cylindrical impedance heads as shown in Fig. 2. One head is used to generate a known acoustic wave at the 'glottis', and a second head at the 'mouth' measures the output wave and determines the pressure and flow at the glottis using only the output signal. The printed vocal tract is fitted between the two heads ($0 \ldots x_M$).

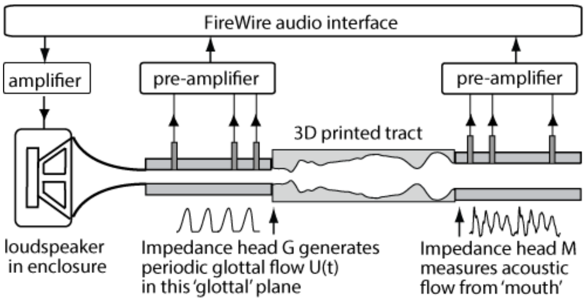

Figure 2. *Sketch of the experimental setup used for generation of known flow and estimation of wave propagation*

In a cylindrical duct, and for wavelengths long compared with the diameter, the formulae that relate pressure and flow at any two cross-sections of the duct are simple and accurate, and results can be improved by calibrating the impedance head using at least two different known acoustic loads (here a very high impedance provided by a 2.4 kg metal mass, and a very long pipe that has very delayed and very attenuated reflections.) The measuring 'mouth' head uses the 3-microphone technique described in [11]. After careful calibration, the measurements from the 3 microphones is used to determine the pressure and flow at x_M (see Fig. 2).

The generating 'glottis' head uses a principle similar to that described in [12]: After calibrating the microphones, the loudspeaker sends an initial glottal flow signal. The resulting signal is measured and after a series of iterations, it approaches the desired acoustic flow at the output $x = 0$. The target and measured flows agree to less than 0.1% as shown in Fig. 3.

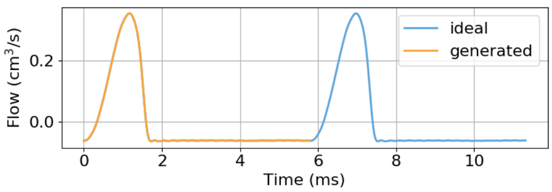

Figure 3: *Target flow pulse (blue) compared to obtained pulse (orange). After iteration, the maximum difference is less than 0.1%*

3. Results

3.1. Benchmark of printing

3.1.1. Accuracy of the print

Using digital photographs of the two halves and by visually recognising the inner edge of the tracts, the printed profiles were compared with the shapes that were provided to the printer and the original profiles. The two latter do not match exactly because the CAD software uses a spline interpolation to fill in the low-resolution original profiles. Fig. 4 shows an example of these measurements on one of the vowel models studied.

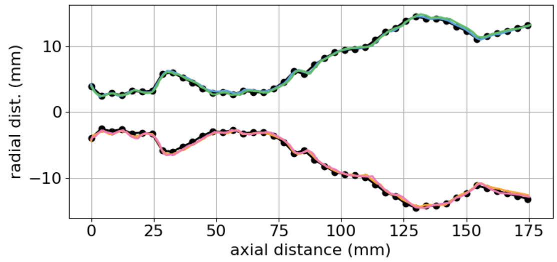

Figure 4: *Measurements of profile from printed two halves compared to the desired geometry*

The two profiles agree to 0.2 mm on average with a maximum difference of 0.5 mm (Fig. 4), an error which is estimated to be smaller than that due to the measurement of the profile in a digital photograph.

Although not clear from figure 1, the 3D printing process adds to the inner surface roughness of about 0.2 mm due to the thickness of the PLA filament. We chose not to polish this to avoid unintended changes to the inner geometry.

Another important factor is the finishing of the two ends of the tract that are used to connect to the impedance heads. The larger base consists of a flat surface that is somewhat polished and to which is applied petroleum jelly to provide a better seal. The opposite end connects to the impedance head via an adapter ring. Due to rugosity, some petroleum jelly is also needed to ensure an airtight seal around this ring.

3.1.2. Acoustic accuracy of printed tracts

The main aim of the 3D prints is to provide an accurate known acoustic system that can be used to test acoustic models. The most basic test is to make sure that a cylinder is accurately printed in order to mimic a cylindrical duct, here a PVC pipe 17cm long cut by hand (Fig. 5). The tract printed in a single piece shows slight changes in positions and magnitudes of resonances which are presumably due to roughness of the walls. After bonding, the 2-piece print has lower frequency resonances than the single piece print although there were no noticeable leaks at DC pressure.

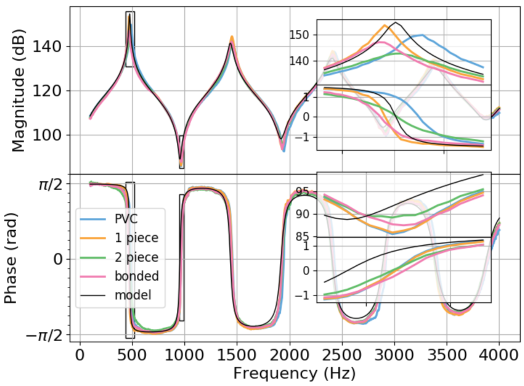

Figure 5: *Comparison of input impedance model and measurements for cylindrical straight ducts (a 17cm PVC pipe, 3D-printed in a single piece, printed in 2 halves before and after bonding). Insets are zooms.*

3.1.3. *Accuracy of the acoustic model of printed vocal tracts*

To test the modelling of changes in area function, a slightly more complex duct with a strong discontinuity was printed: it consists of a 170 × 7.8 mm diameter tube connected to a 26.2 mm tube, also 170 mm long. Figure 6 shows that the positions of secondary resonances (splitting of the full-length resonance due to discontinuity in cross-section) are not well captured by the model, indicating the need for a better model of acoustic impedance when discontinuities occur in the area function $A(x)$.

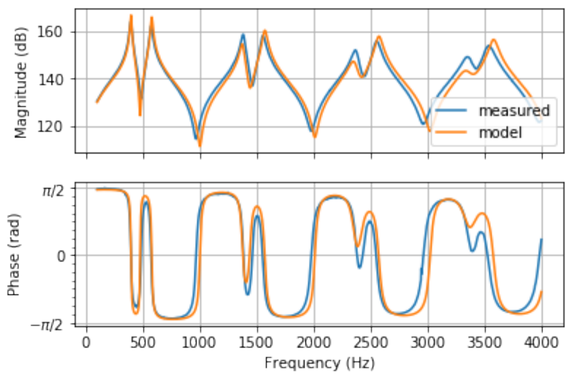

Figure 6: *Acoustic impedance of a cylinder 7.8 mm in diameter stepping into a cylinder 26.2 mm in diameter, compared with the model.*

1d computational models and 3D printed ducts using area functions from [10] usually fit well for lower frequencies (up to around 1kHz) with larger deviations in the 1-4 kHz range. One example, for vowel /a/, is shown measured from the glottis with an open mouth (Figure 7) and from the mouth with a closed glottis in Figure 8. The differences are usually seen in the Q factor of the resonances and seem more important when there are more sudden jumps in the area function.

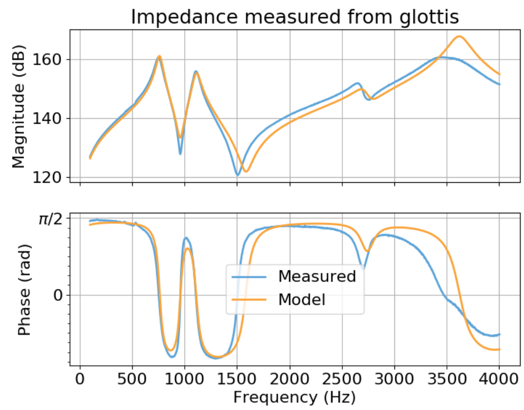

Figure 7: *The impedance measured on a 3D printed tract. Z at 'glottis' with an open 'mouth'; it is compared with a 1d model calculated with 100 segments*

3.2. Estimation of acoustic variables

In a later experiment, the signals recorded by microphones in the 'mouth' head were used to determine pressure and flow at x_M, and from these, using models of the propagation inside the tract with known geometry, the pressure and flow were estimated at $x = 0$.

For the vocal tract shown in Figures 1 and 4, the target flow and the estimated flow are shown in Figure 9, from which it is apparent that the lower frequency components are reasonably well estimated, whereas the higher frequency ones are much less accurate.

The agreement is fairly good for measurements at the fundamental frequency of the speaking voice, but by 700 Hz it is of the order of 3 dB with a phase difference of order $\pi/4$. Above 1 kHz the estimation is poorer with errors larger than 10 dB and phases that are almost uncorrelated with the target signal. These surprisingly large differences may be due to relatively rapid variations in $A(x)$ and thus the combination of multiple errors of the type seen in Figure 6. The poor performance at high frequency means that the technique is currently of little use for phonetics. However, for understanding the auto-oscillation of the folds and their interaction with acoustic loads at the fundamental frequency, it may have promise because it makes few strong assumptions.

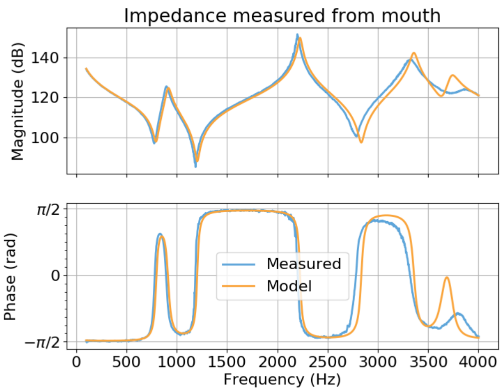

Figure 8: *As for Figure 7, but this time measured at the 'mouth' with a closed 'glottis'*

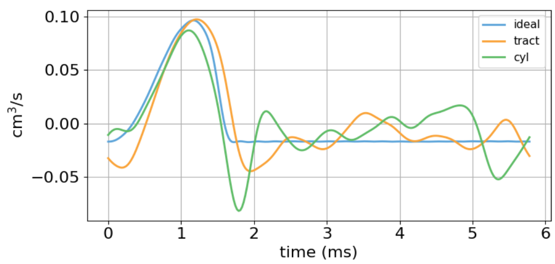

Figure 9: *Target flow pulse (blue) and estimated pulse from measurements at the mouth using propagation models based on the tract geometry (orange) or simply approximating it as a straight cylinder (green)*

3.3. Measuring downstream impedance from a glide

Given that the main purpose of this work is to deduce the acoustic wave generated by the vocal folds, the system was also tested in real-life conditions. It is the long-term aim to replace all the experimental apparatus upstream of the measuring impedance head with the mouth of a subject, where the $A(x)$ could be determined by the method described in the companion paper [13]. At the end of the mouth impedance head, a loudspeaker is fitted that is used for a different experiment. In this preliminary experiment, it is shown that the human voice may be used to measure the impedance of the system downstream from the mouth.

The hardware setup of the impedance head is similar to the one in the previous experiment, however the analysis of the signals recorded by microphones 1-3M is slightly different, because the excitation signal is no longer strictly periodic.

If the singer performs a pitch glide that extends over at least one octave, there is, in principle, enough information in the microphone signals to extract the complex amplitudes at each frequency. From these, and given the same calibration data as for the previous measurement, it is possible to calculate pressure and flow of the plane wave travelling inside the impedance head. The ratio of pressure and flow gives the acoustic impedance of the passive system downstream of the head (Fig. 10). This information is important to determine the load acting on the vocal folds.

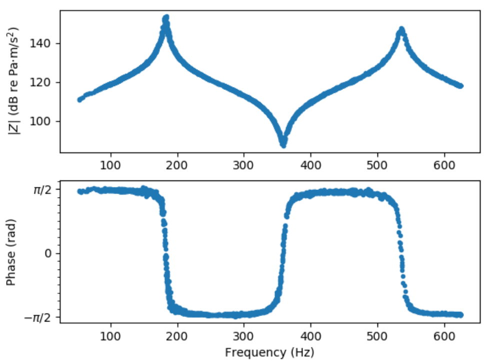

Figure 10: *Downstream impedance of one of the measurement heads open at the driver end, calculated from the first harmonic of a sung pitch glide*

4. Discussion

3D printing provides a quick and affordable way of producing test geometries that reproduce approximately the acoustic loads seen by the vocal folds at frequencies in the range of the fundamental f_0 of speech.

In order to produce 3D printed tracts that are useful for acoustic measurements, the ducts must be printed in a way allows quality control of the geometry. In particular, extraneous filaments should be removed, and the geometry of the print should be checked against the desired geometry. The disadvantage of this process is that the duct must be printed in two or more pieces and then glued together after polishing the contacting surfaces. Extra care needs to be given to ensure the airtightness of the two halves after gluing and the fit of the duct to the measurement device, for example by using grease to seal the contact. Air leaks decrease the magnitude of measured impedance peaks.

In these conditions, the acoustic properties below 1 kHz are relatively well predicted by a simple propagation theory that considers 1 dimensional plane waves (*e.g.* Figs 6-8). Geometries with rapidly varying cross-sections are particularly challenging and will probably need a more sophisticated propagation model. Using the current propagation model, calculations that treat the (complicated) duct as a simple cylinder provide results whose accuracy compares with that calculated using the real geometry. However, simpler experimental geometries such as straight ducts and slowly varying cross-sections provide an extended range of usable frequencies up to 4 kHz, when calculated with the real geometry. These results point towards improving the propagation model at discontinuities using lumped elements that include an inertive end effect and a stagnant air compliance.

The calculations of the transfer matrices can be compared with the use of inverse filtering, e.g. in [4]. In that work, the errors in the estimation of the glottal flow are significantly smaller than in the current study, although in that study the injected flow follows very closely the assumptions of the inversion method, in particular the frequency content of the flow signal follows the model used in the inversion method.

5. Acknowledgements

We thank the Australian Research Grants Council for support.

6. References

[1] Rothenberg, Martin. “A new inverse-filtering technique for deriving the glottal air flow waveform during voicing.” The Journal of the Acoustical Society of America 53(6): 1632-1645, 1973.

[2] Alipour, F., and Scherer, R. C., “Pulsatile airflow during phonation: an excised larynx model” J. Acoust. Soc. America, 97(2): 1241-1248, 1995.

[3] Airas, M., “TKK Aparat: An environment for voice inverse filtering and Parameterization,” Logopedics Phonatrics 33(1): 49–64, 2008.

[4] Chu, D.T.W., Li, K., Epps, J., Smith, J. and Wolfe, J., “Experimental evaluation of inverse filtering using physical systems with known glottal flow and tract characteristics” J. Acoust. Soc. America. 133(5): EL358-362, 2013.

[5] Alku, P. “Glottal wave analysis with pitch synchronous iterative adaptive inverse filtering.” Speech communication, 11(2-3), 109-118, 1992

[6] Kitamura T, Takemoto H, Adachi S, Honda K. “Transfer functions of solid vocal-tract models constructed from ATR MRI database of Japanese vowel production.” Acoust. Sci. Tech. Tech. 30(4):288–296, 2009.

[7] Delvaux B., and Howard D. “A New Method to Explore the Spectral Impact of the Piriform Fossae on the Singing Voice: Benchmarking Using MRI-Based 3D-Printed Vocal Tracts,” PLOS ONE 9(7): e102680, 2014. https://doi.org/10.1371/journal.pone.0102680

[8] Echternach M., Birkholz P., Traser L., Flügge T. V., Kamberger R., Burk F., et al. “Articulation and vocal tract acoustics at soprano subject’s high fundamental frequencies.” J. Acoust. Soc. America 137(5):2586–2595, 2015. pmid:25994691.

[9] Fleischer, M., Mainka, A., Kürbis, S. and Birkholz, P. “How to precisely measure the volume velocity transfer function of physical vocal tract models by external excitation” PLOS ONE 13(3): e0193708, 2018.

[10] Story, B. H., Titze, I. R., & Hoffman, E. A. “Vocal tract area functions from magnetic resonance imaging.” J. Acoust. Soc. America, 100(1): 537-554, 1996.

[11] Dickens, P., Smith, J. and Wolfe, J. (2007) “Improved precision measurements of acoustic impedance spectra using resonance-free calibration loads and controlled error distribution”, J. Acoust. Soc. America, 121(3): 1471-1481.

[12] Wolfe, J., Chu, D., Chen, J.-M. and Smith J., “An experimentally measured Source-Filter model: glottal flow, vocal tract gain and radiated sound from a physical model”, Acoust. Australia, 44: 187–191, 2016.

[13] Hanna, N., Rodriguez, A., Almeida, A., Wolfe, J., Smith, J., “Estimation of vocal tract and trachea area functions from impedance spectra measured through the lips”, Speech Science and Technology Conference, 2018