Adjustment of the Vocal Tract Shape via Biofeedback: A Case Study

*Patrick Hoyer, and †Simone Graf, *†*Munich, Germany*

Summary: In this study, an adjustment of the vocal tract shape toward selected sound waves in the frequency range of the first and second formants without phonation is discussed. The sound waves of a loudspeaker in front of the open mouth and amplified by the vocal tract are used as biofeedback signals. It is shown that the resonance amplification of the vocal tract complies with the concept of forced oscillation, with the driver being the sound source and the resonator being the vocal tract. An adjustment toward increased amplification via vocal tract resonance can be related to smaller bandwidths and lower damping. Furthermore, the applied adjustment frequencies are preserved as vocal tract resonances during exhalation and even phonation. This novel form of biofeedback might enrich standard voice training procedures by exercises without phonation.

Key Words: Biofeedback−Vocal tract−Formants−Bandwidths−Voice training.

INTRODUCTION

The human voice is a fundamental element for social interaction and self-expression, and the combination of well-defined vowels with consonants is the base for any human language. Successful voice development relies on learning new—or relearning previous—vocal motor behaviors which result in optimized output qualities (eg, loudness, pitch, voice quality, respiration, efficiency).[1,2] Voice production, especially during singing, is a complex process where different aspects such as the stance,[3] breathing, glottal activity,[4] and shaping of the vocal tract are simultaneously addressed.[5,6]

The sound of a healthy voice can easily be recognized; however, it is desirable to rely on clearly defined objective measurement procedures to monitor the state of the voice and its development for each of the variables independently. There is a large variety of analytical methods that can be used for respiratory analysis or posture, on the one hand, and of the voice during phonation, on the other hand. In regard to feedback procedures on phonation, microphone measurements are widely used to improve pitch and loudness of the singer or speaker with short response time.[7]

Among other elements essential for voice production, the highly flexible vocal tract is the key element for vowel production. An adjustment of its resonances (formants) is possible by means of the articulatory elements (ie, lips, tongue, jaw, and larynx).[8]

The present work discusses a possible fast biofeedback mechanism for an adjustment of the vocal tract shape toward specific frequencies within the range of the first and second formants. For this purpose, an external sound field in front of the mouth is offered and the participant is asked to actively adapt the articulatory elements of the vocal tract to yield a maximum amplification without phonation (ie, without oscillation of the vocal folds).

In the literature, there are several approaches to measure acoustic data of the vocal tract. Fujimura and Lindquist excited the vocal tract via the neck obtaining the information about the transfer function.[9] The reverse measurement is also possible using a three-dimensional model of the vocal tract derived from MRT studies.[10] In this case, the plastic model is loaded by an external sound source at distance and a microphone located at the position of the glottis.

Rothenberg used a tube close to the mouth driven by a membrane to stimulate vibrations across the vocal tract at the vocal folds.[11] Highly accurate impedance data in a broad wavelength spectrum could be realized via an impedance tube even during phonation.[12] Epps et al[13] introduced a method to measure the load of the vocal tract with an acoustical duct close to the mouth and a microphone at a similar position. In this arrangement, the impedance of the resonator (vocal tract) can be separated from the output impedance of the sound source (loudspeaker) by introducing a high resistive load within the acoustic duct, limiting the effects of tube resonances.

In our study, a sound source close to the mouth and similar to the ones used in mobile phones interacts directly with the vocal tract of the participant at specific frequencies. The sum of the acoustic signals of the external sound source and the amplification due to the resonance of the vocal tract are used as biofeedback signal. Furthermore, the effect of the amplification on the acoustic properties of the vocal tract is discussed.

EXPERIMENTAL PROCEDURE

Instrumentation

Acoustical measurements have been conducted within an audio booth of the phonetics department of the TU Munich or the Fraunhofer-Institute of Integrated Circuits in Erlangen, the latter with a residual sound level of less than 20 dB and spectrally linear reverberation. A Schoeps (Germany) MK2 microphone with a preamplifier CMC 6 was used,

Accepted for publication January 19, 2018.
From the *Fraunhofer Headquarters, Munich, Germany; and the †Department of Otorhinolaryngology/Phoniatrics, Klinikum rechts der Isar, Technical University Munich, Munich, Germany.
Address correspondence and reprint requests to Patrick Hoyer, Fraunhofer Headquarters, Hansastr. 27c, 80686 Munich, Germany. E-mail: patrick.hoyer@zv.fraunhofer.de
Journal of Voice, Vol. 33, No. 4, pp. 482−489
0892-1997

https://doi.org/10.1016/j.jvoice.2018.01.018

which was confirmed by a GRAS 46AF (G.R.A.S Ltd. UK) microphone connected via a preamplifier Nexus 2690 (Brüel&Kjaer, Denmark) to a Steinberg UR242 (Yahama Inc., Japan) as an external sound card. The distance between the microphone and the vocal source was 20 cm. All recordings were performed at 48,000 Hz/16 bit using the software *Audacity* (D. Mazzoni, Mountain View, CA, USA).

Spectrograms and Fast Fourier Transform (FFT) spectra were obtained using the *WaveSurfer* (Sjölander and Beskow, Stockholm, Sweden) Version 1.8.8.p4 software. The FFT analysis during phonation was performed at wavelength steps of 24 Hz (2048 FFT points), spectrogram of the exhaled air and under phonation was analyzed with an analysis bandwidth of 4 Hz.

The sound source for training and amplitude measurements shown in Figure 2 was an inbuilt buzzer of an iPhone 5s (Apple inc., Cupertino, CA, USA) with single frequencies supplied by a frequency generator app (ee-toolkit) whose accuracy was measured with a lock-in amplifier (Stanford Instruments 830 DSP) to be less than 0.1 Hz of error. The distance between microphone and mouth was 20 cm.

Phase and amplitude of the resonating wave were detected with a lock-in amplifier Stanford Instruments model 830 DSP (Stanford Instruments, Sunnyvale, CA, USA), and the data of amplitude and phase were recorded with a Picoscope 2000 (Pico Technology, Cambridgeshire, UK) at 12-bit resolution. The time constant of the lock-in amplifier was set to 10 milliseconds throughout the experiments. The sound source used for the phase-related measurements was an external iPhone 5s replacement buzzer with an opening of 2×9 mm^2 connected to an iPhone 5s using the frequency generator app (ee-toolkit) as stated previously with one channel connected to the lock-in amplifier and the other to the buzzer. However, similar results were obtained with a Visaton BF 32 (Visaton, Hahn, Germany) speaker in a sealed chamber connected to a 5-cm tube (6 mm in diameter). The microphone was placed at a distance of 20 cm from the mouth. Zero-degree phase shift refers to the condition without resonances (closed-mouth position).

Procedure

The participant adjusted the vocal tract toward external acoustical sinusoidal sounds in the range of 400−2000 Hz. A build-in buzzer of an iPhone (Apple inc., Cupertino, CA, USA) 5s was used as the sound source, which was applied close to the open mouth and restricted to 60 dB sound pressure level at 20 cm.

During a training period of 2 months, the task given was to enhance the acoustic signal by adjustment of mouth, tongue, jaw, throat, neck, etc., without phonation. The optimization was achieved via acoustical biofeedback without additional signal processing. Thus, the total sound pressure of the source in connection with the resonant vocal tract was used as biofeedback signal. After the training period, the participant was able to sustain the position of highest resonance comfortably for several seconds, whereas the resonances adjusted to frequencies between 700 and 1300 Hz were most stable.

Before each of the presented measurements, the participant enhanced the amplitude of the external sound source by variation of the vocal tract without phonation 10 times for 30 seconds within 15 minutes to predefined wavelengths (adjustment frequencies) of the 700, 900, and 1300 sinusoidal tones, respectively (adjustment period). Note that all adjustments of the vocal tract toward the offered frequencies were performed while breathing through the open mouth and without using the vocal cords (without phonation).

After the adjustment period, the participant was asked to exhale air or phonate in the position found during the adjustment period. First, exhalation was repeated for at least three times and—after further 30 seconds of adjustment—the phonated sound was repeated as at least three times. All sounds were sustained for at least 3 seconds and the participant was asked to take a breath between each sound production. Vocal tract adjustment, exhaling, and phonation were repeated at each adjustment frequency several times with similar results.

During the phase-related experiments used for Figure 3, the participant was asked not only to enhance the signal by adjustment of the vocal tract but also to vary the vocal tract position to pass periodically through the adjustment position of highest amplification.

All experiments were conducted with one participant with a healthy voice (nonprofessional, choral experience [bass] > 10 years), and an average fundamental frequency in speech: 99 Hz (44 dB), sound pressure level maximum: 104 dB (A), measured with a XION Voice Analysis System (XION GmbH, Berlin, Germany). The experiments are approved to be compliant with ethical standards set by the Ethics Committee of the Technical University Munich in January 2017.

THEORY AND BACKGROUND

In the present study, the acoustic parameters of the human vocal tract are altered deliberately to meet the resonance conditions of an external sound source at a fixed wavelength.

In the concept of forced oscillation, a maximum amplification will be met if the frequency of the driver meets the resonance frequency f_0. Furthermore, damping will broaden the resonance in respect to ω and reduce the amplitude. The amplification (resonance) of the resonator (ie, the vocal tract) is thus a synonym for low bandwidth (B_w) or high quality factor (Q) at the given frequency. In this case, the maximum amplification A will be limited by the loss factor α by

$$A_{max} \sim \frac{1}{\alpha} \sim \frac{f_0}{B_w} = Q. \quad (1)$$

The quality factor may also be interpreted as energy stored divided by the energy dissipated. It is obvious that high amplitudes will be obtained at the resonance frequency and minimized loss factor.

The phase lag φ between acoustical driver and the amplified sound of the resonator is determined by

$$\tan\varphi = \frac{2\alpha\omega}{{\omega_0}^2 - \omega^2}, \tag{2}$$

with ω being the angular frequency $2\pi f$ of the driver. Thus, a phase shift of $-\pi/2$ is observed if frequencies of driver and resonator are equal. The phase shift is zero either with no coupling between resonator and driver or the resonator is tuned to a significant higher frequency. A phase shift of $-\pi$ will be found at significantly lower resonance frequencies of the resonator compared with the driver frequency.

The vocal tract is regarded as an open-closed tube which resonance frequency is tuned to an externally offered sinusoidal signal. The resonance conditions of an open-closed tube are determined by the radiation impedance at the tube opening, which can be described with the assumption of $\alpha L \ll 1; \frac{\alpha}{k} \ll 1$ by

$$\frac{Z_{m0}}{\rho_0 cS} = \frac{\alpha L - i\cos kL \sin kL}{\sin^2 kL + (\alpha L)^2 \cos^2 kL}, \tag{3}$$

where Z_{m0} is the radiation impedance at the tube opening, L is the length of the closed tube including an end correction, $k = 2\pi/\lambda$, and α is the loss coefficient.

The complex and the real parts of the impedance at the tube opening are high at $kL = n * \pi$ and the radiation is small (antiresonances). At the resonance frequency, the imaginary part of the impedance will be zero at $kL = \left(2n + \frac{1}{2}\right) * \pi$, with a pressure node at the open end and a pressure antinode at the closed end. In relation to the particle velocity, an antinode is formed at the open and a node at the closed end. With the imaginary part being zero at resonance, the impedance is reduced to its real part and the amplitude is determined by the loss factor α. The real part of Eq. (3) at resonance conditions may thus be related to the sum of different losses (ie, reflection losses, wall interactions, radiation, and viscous or thermal absorption).

As stated previously, a driver with its radiation impedance Z_{md} at the tube opening can excite the acoustic system toward oscillation. In order not to alter the resonance conditions of the vocal tract, the impedance of the driver should be low, which will hold true for acoustic coupling via air and with a high particle velocity at the interface between the driver and the open mouth.

Coupling will be most effective if the driving sound source is located within the end correction term of a tube opening, that is, within a distance of about 0.8*r, with r being the radius of the mouth opening.[14] In the present study, this was realized by using a speaker close to the mouth. In addition, resonances of any waveguide can be neglected if the distance of the oscillating membrane of the speaker to the open mouth is chosen to be within the lumped region at the wavelengths used.

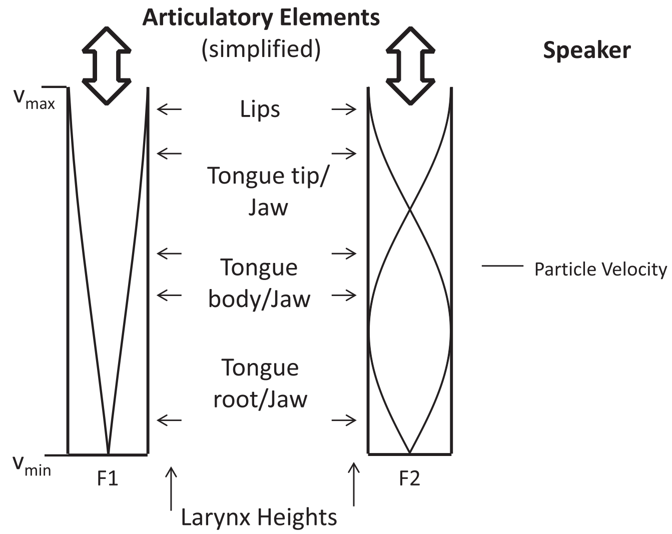

FIGURE 1. Simplified model of the excitation of the vocal tract with an external sound source: resonance condition of the first and second formats (F1, F2) of the vocal tract. Basic articulatory elements to tune the vocal tract resonances are indicated. Adapted from Lidblom and Sundberg and Wolfe et al.[8,15]

Figure 1 illustrates the arrangement at resonance condition in respect to the particle velocity for the first and the second formants. Note that the acoustic pressure at resonance (not shown in Figure 1) demands a node at the mouth and an antinode close to the glottis within the acoustic tube model. Under this condition, the sound waves are reflected from the open and the closed side, forming the so-called standing waves, which is equal to the concept of constructive interference of waves.

RESULTS

Figure 2 shows the changes in the sound level during the adjustment of the vocal tract shape toward frequencies of 700, 900, and 1300 Hz, respectively. Alterations within the articulatory shape of the vocal tract result in strong changes of the overall sound level, and the participant needed some time to find the position of the highest amplification. The procedure of adjusting the vocal tract shape, as shown in Figure 2, was repeated until a stable position could be reproduced (adjustment period).

Using a lock-in amplifier, the phase and amplitude of the resulting sound wave can be detected simultaneously. Apart from an increase in amplitude, the phase of the signal used as biofeedback is shifted as shown in Figure 3.

The sound pressure amplitude is clearly linked to the phase shift between the driving source and the amplification due to the resonance of the vocal tract. In all cases, the highest amplitude may be found at a phase shift of the resulting wave at about $-\pi/2$. In the case of the resonance amplification at 1300 Hz, the formation of antiresonance at the same phase shift and minimum of sound pressure is observed as well. In the following table,

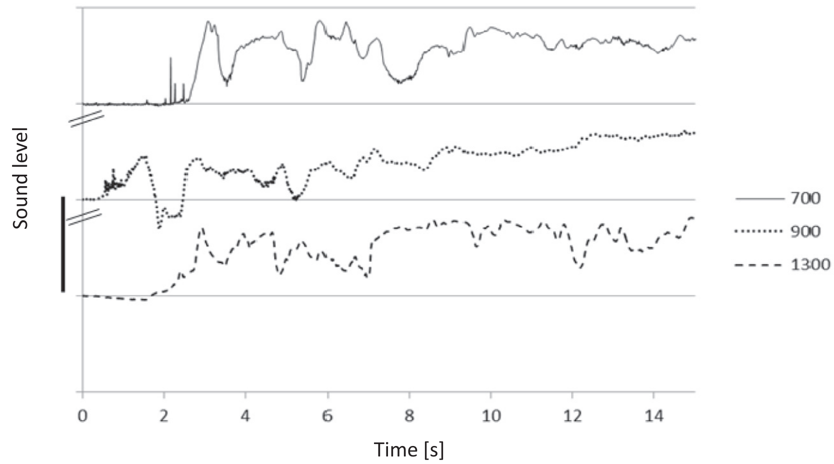

FIGURE 2. Changes of the sound level of the biofeedback signal during the adjustment of the vocal tract shape to resonance at frequencies of 700, 900, and 1300 Hz. The bar and the spacing represent a difference in the sound level of 20 dB. The baseline represents the signal with the mouth closed at the respective frequencies.

typical values of resonance amplification are listed (antiresonances are omitted).

	700 Hz	900 Hz	1300 Hz
Amplification at −90° phase shift (dB)	8–16	7–15	6–17
Phase shift at highest amplitude (°)	−(80–100)	−(80–100)	−(80–100)

It is evident that an adjustment of the vocal tract toward one frequency will affect the acoustical environment of the voice production system as a whole. To observe changes in the vocal tract without using the vocal folds, the resonance spectrum of an exhaled air might be a suitable indicator.[16] Figure 4 shows sound pressure and spectrogram of the exhaled air after adaptation of the vocal tract. Note that no vowel has been intentionally formed and the position of the highest resonance intensity was asked to be maintained.

After the adjustment period, a stable position could be sustained for several seconds and resonances are observed at the adjustment frequencies which could be reproduced within 10% of the frequency applied. The maxima of what might be attributed to the first formant shifts from around 700 to about 450 Hz as the adjustment frequencies increase. In addition, a higher resonance is observed to shift from around 2400 Hz (adjustment frequency: 700 Hz) to lower values.

Figure 5 shows the sound pressure and spectrogram during phonation after the adjustment periods. Although there is no intentional articulation of any vowel, the sound after adjustment toward 700 Hz is similar to an “o,” whereas the sound after 900 Hz has an appearance of an “a” with a tendency toward “ae” after 1300 Hz.

The basic features of the spectra are found similarly during phonation and exhalation, with the adjustment frequency being supported. In the spectrogram, the higher and lower regions are shifted with increasing adjustment frequencies, that is, after an adjustment of the vocal tract adjustment toward frequencies of 1300 Hz, the lower and higher resonance are found at about 500 and 2400 Hz, respectively.

Figure 6 shows the FFT spectra of the waveforms shown in Figure 5.

The resonance tuning can clearly be observed in the harmonics during phonation in the frequencies at 700, 900, and 1300 Hz. In addition, the highest harmonics of intensity are reduced from H4 to H3 and from H20 to H17 as the adjustment frequency (before phonation) is increased.

DISCUSSION

Adjustment of the vocal tract shape and resonance

The participant used the audible resonance of the vocal tract excited by an external sound source as biofeedback to adjust the articulatory elements toward a stable position without phonation.

In the model of a forced oscillation (Figure 1), the external speaker is considered as driver and the vocal tract as resonator. The most intense coupling will demand that (1) the resonance frequency is met and (2) the bandwidth of the resonator is small, the latter being related to a small loss coefficient. In the experimental procedure used, the driving frequency is kept constant, whereas the vocal tract of the participant (resonator) is adjusted to meet resonance conditions. As the participant was asked to enhance the amplitude of the external sound source (Figure 2), the most apt condition can be found within the first 10 seconds. In the case of the adjustment toward 900 Hz, a reduction of the overall amplitude at 2 seconds is observed. In this case, the vocal tract seems to have been tuned shortly to an antiresonance at 900 Hz.

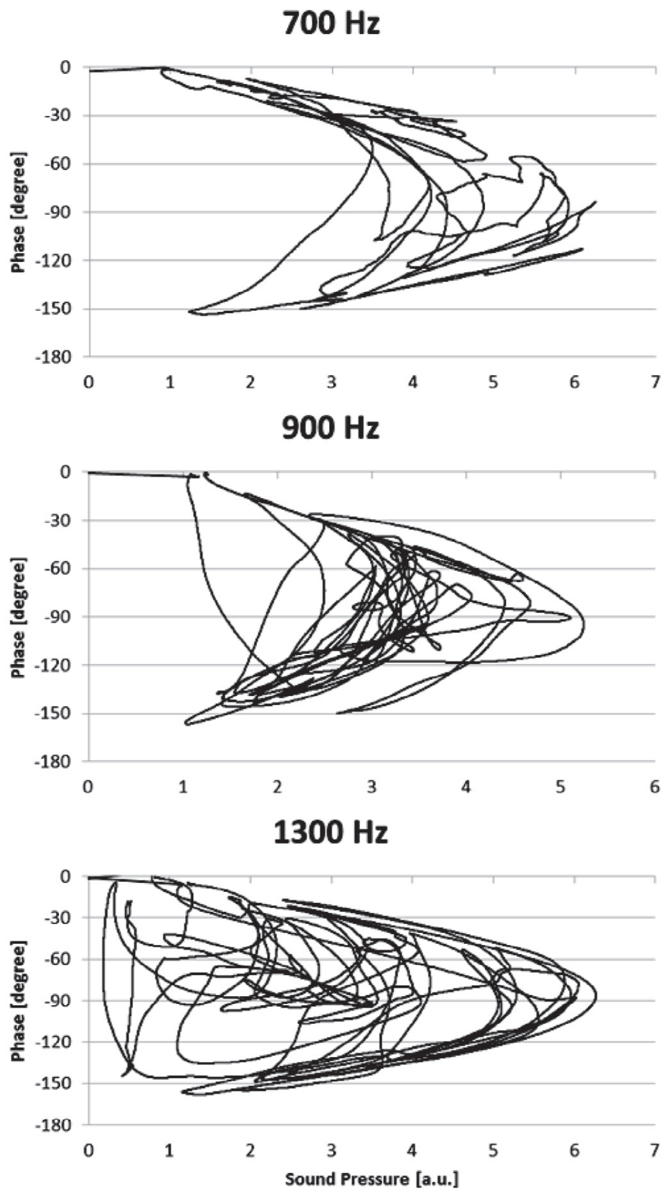

FIGURE 3. Phase shift and amplitude of the biofeedback signal during adjustment of the vocal tract shape toward the frequencies indicated. A zero-phase shift refers to the condition without resonance (closed mouth). Note that the sound pressure amplitudes are plotted linearly.

With the mouth closed, only the amplitude of the driver is measured. In the case of the opened mouth, coupling occurs via air and the additional sound from the oscillator adds to the one of the driver. As shown in Figure 3, a phase shift is observed if the vocal tract is shaped toward resonance at the given frequency. A maximum amplitude is found at a phase lag of $-\pi/2$ (90°) between the loudspeaker and the resonator. From Eqs. (2) and (3), a phase lag other than $-\pi/2$ may be directly linked to the difference between the driving frequency of the speaker and the resonance frequency of the vocal tract.

As the lock-in amplifier used in Figure 3 records the sum of the acoustic signals—the one of the driver and the resonant signal of the vocal tract simultaneously—the phases of the waves add as well. At maximum amplification between 16 and 18 dB in relation to the acoustic signal of driving source alone (closed mouth), the observed shift is reduced by only 10° to −80° but will be higher at low amplification. This would be in accordance with the values in Figure 3 at lower intensities between $-\pi/2$ and $-\pi$. However, the phase lag of less than −90° may also be influenced by the presence of other resonances (ie, formants) in the vicinity of the adjustment frequencies.

However, the amplitude depends not only on the resonance matching but also on the loss coefficient α. Thus, besides changing the resonance condition of the vocal tract to meet the adjustment frequency, high amplitudes at a phase lag of $-\pi/2$ are interpreted as a synonym for a low loss coefficient (ie, a small bandwidth, a high quality factor) (Eq. 1). It is evident that a low loss coefficient will be helpful for facile voice production, which may be supported by using the proposed biofeedback.

From the experiments, it can be deduced that the loss coefficient of the vocal tract may be reduced by vocal tract adjustment without phonation. However, the measured results do not specify the origin of the effects, leading to reduced damping. In a tube-like system, viscous and thermal losses are usually low. Main losses occur via contact of the oscillation air with the mucous membranes, the radiation of sound, the shape of the vocal tract, and the reflectivity at the closed end.[8] Although the tissue is regarded as acoustically reflective at frequencies higher than 500 Hz,[1] the radiated power will increase with frequency. Given a constant wavelength of the signal and only small variations of the mouth opening, radiation losses may be considered as stable. Thus, it is reasonable to assume that losses due to wall interactions or radiation during resonance may only be addressed within a narrow region by the adjustment. Consequently, the shape of the vocal tract and a high reflectivity especially at the pressure nodes or antinodes will be discussed in more detail.

Equation (3) demands a pressure node at the mouth and a high impedance (reflection) condition at the closed end for the evolution of stable "standing" waves (ie, resonance). Although a pressure node close to the mouth can be assumed to prevail as a resonance condition and the location where the sound waves are reflected, the acoustical boundary conditions inside the vocal tract are less obvious. The basic physiological structure of the vocal tract may give a hint as where to find possible (and tunable) zones for high amplification of the induced resonances. In the literature, there is strong evidence that during phonation, as well as during whispered speech, an antipressure node is located at the glottis.[10,17] However, there might be more flexibility to meet this condition without phonation. As all experiments took place with an unrestricted airflow from the trachea through the lips, the vocal folds may not be regarded as a high impedance barrier. The congruence of the resulting spectra after resonance adjustment by exhaling air and during phonation at frequencies of 700−1300 Hz, however,

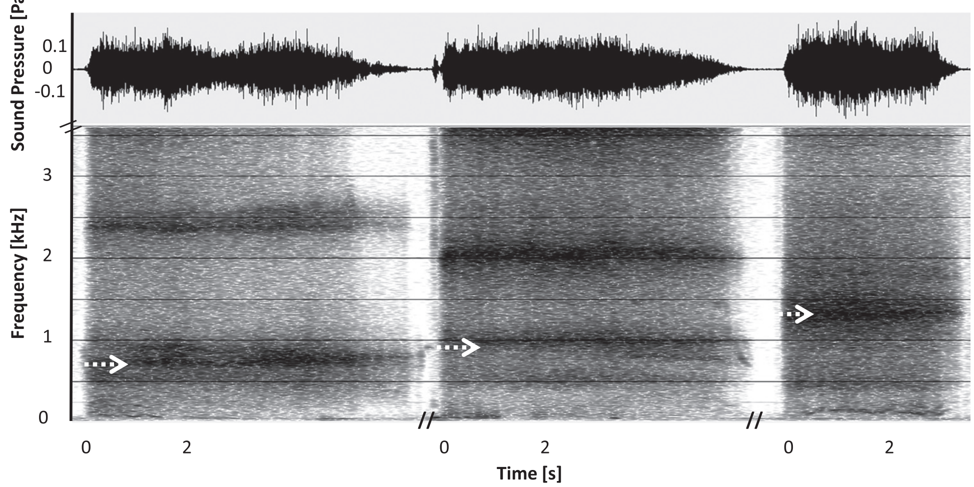

FIGURE 4. Sound pressure and spectrogram of the exhaled air after an adjustment period using frequencies of 700, 900, and 1300 Hz, respectively. Top: sound pressure of the exhaled air; bottom: spectrogram from 0 to 3600 Hz of the exhaled air, with the *arrows* indicating the adjustment frequencies positions. Darker areas represent higher intensities.

makes it reasonable to look for anatomic structures that will be able to represent a tunable high impedance reflection barrier not far from the high impedance resonance condition present during phonation. Hanna et al. studied the acoustic properties during phonation in comparison with the glottis closed[18] and found little changes between the two positions in regard to the transfer function of the vocal tract. The epilarynx tube with its small area function compared with the pharynx has been discussed as a possible barrier and high impedance source for a conversion of the airflow through the glottis toward high pressure at the bottom of the vocal tract.[17] In this context, the high impedance at the transition of the pharynx to the epilarynx tube may act as a restriction with high reflectivity and thus the proposed pressure antinode. Furthermore, if the high impedance necessary for resonance condition is found in this region, an adjustment of the acoustical behavior toward external frequencies should have a positive impact on the coupling of the vibration of the vocal folds with the vocal tract as well.[19] In this context, it should be noted that during tube phonation, a phase lag between the driver (vocal folds) and the vocal tract of $-\pi/2$ at maximum coupling has been discussed.[20]

The proposed driver-resonator-model related to Eqs. (1) and (2) is not linked to any specific geometry of the vocal tract. Apart from the previously discussed tube model, where adjustment of the vocal tract shape is directly related to the length of the resonator, low bandwidth resonance may also be achieved in Helmholtz resonators. In a Helmholtz resonator, the volume in a small neck represents the mass of the vibrating system and is connected to a volume

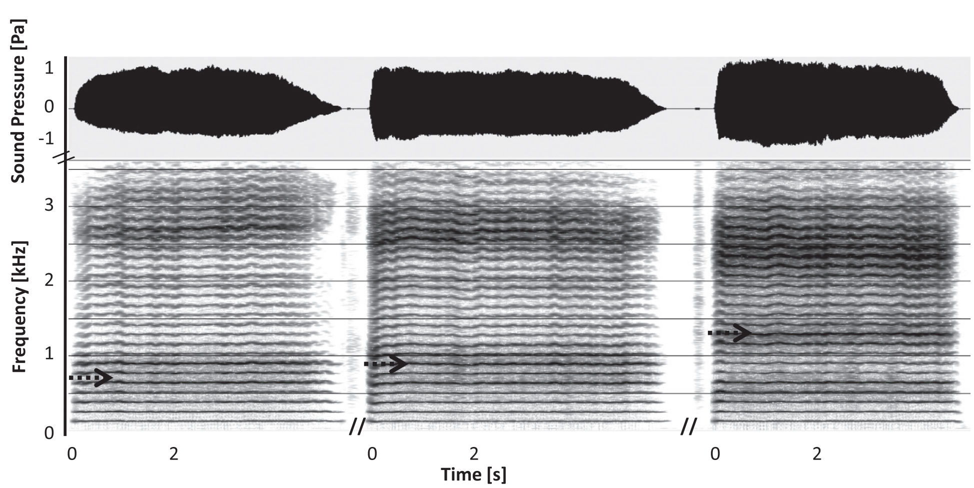

FIGURE 5. Phonation at a fundamental of $f = 131$ Hz after adjustment period of 15 minutes at frequencies of 700, 900, and 1300 Hz, respectively. Top: sound pressure of the phonated sound without active vowel formation; bottom: spectrogram of the phonated sound with the adjustment frequencies position indicated by *arrows* at 700, 900, and 1300 Hz, respectively. The plot shows the spectrum from 0 Hz to 3600 Hz. Darker areas represent higher intensities.

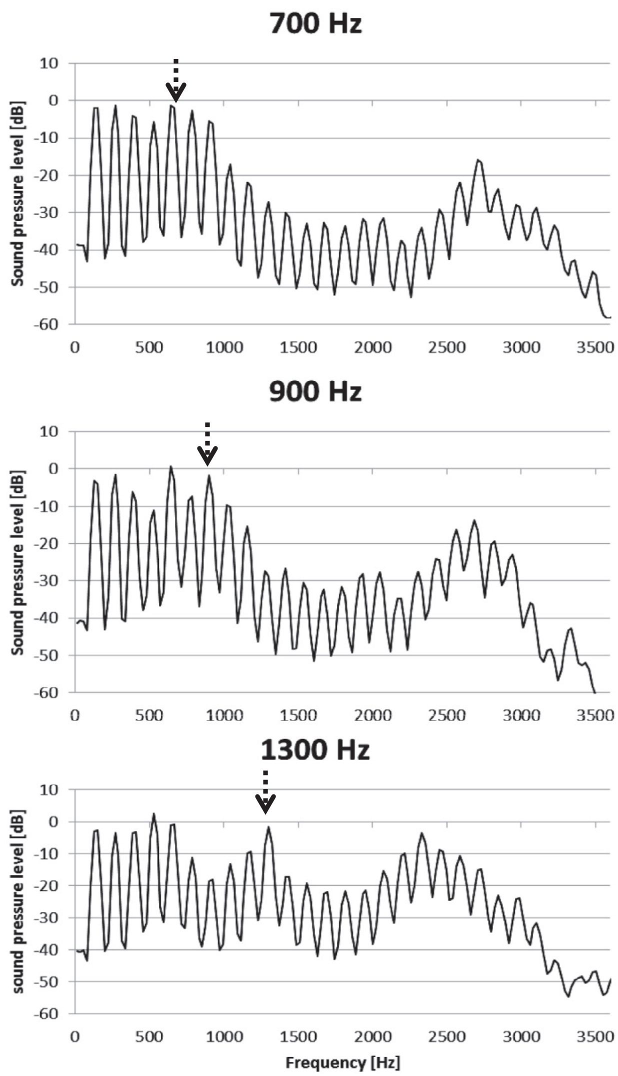

FIGURE 6. FFT analysis after the adjustment period (f_0 = 131 Hz). The analysis is a 3-second average of a selection taken at the center of the produced sounds after adjustment of the vocal tract shape to 700, 900, and 1300 Hz (Figure 5). The *arrows* indicate the adjustment frequency.

of spherical size, acting as a spring. Also, in this model, a pressure node is present at the mouth opening. The transition between a tube model and a Helmholtz resonator within a model pipe system with two different diameters occurs gradually.[21] This more “Helmholtz-like” position may be realized with a large area function at the pharynx compared with the area function of the mouth.

Vocal tract adjustment and formants

Human vowels are identified by the position of the first and second formants, which can be interpreted as vocal tract resonances (Figure 1). As the concept of format formation is rather related to the shape of the vocal tract than to vowel recognition, we call the observed increases in intensity “formants” in the following discussion.

During the adjustment of the vocal tract, only one frequency is actively amplified. However, other resonances are affected during the positioning, also, as can be seen in the spectra of the exhaled voice and during phonation. Note that no intentional change in the vocal tract position was made and the adjustment took place at normal breathing.

Studies showed rising formants in the whispered voice compared with phonation,[4,22] and the changes were explained by a reduction in the diameter of the vocal tract close to the glottis. In Figure 3, the transfer function at open glottis condition during exhaled air is close to the position during adjustment of the vocal tract shape. Exhaled air seems to preserve the vocal tract resonances, whereas whispering may be regarded as being related to intentional changes within a separate communication mode. It is surprising, however, that the frequencies used for adjustment find their representation during phonation as well. Obviously, there must be changes from the open vocal folds toward a position supporting phonation with the folds partially closed. However, we can regard resonance as the condition where sound waves reflected from the closed and the open end result in a constructive interference (Figure 1, “standing” waves). In this concept, the location of excitation of the resonator itself will not change the resonance frequencies as long as the excitation occurs compatible with the resonance conditions. With the excitation of the vocal tract during the adjustment phase with high particle velocity at the open mouth, this condition should be fulfilled. The proposed mechanism should be supported by the high quality factor of the vocal tract induced during the adjustment phase.

After an adjustment frequency of 700 Hz, the first and the second formants are found close to each other (phonated sound similar to an “o”). As the adjustment frequency is increased, the second formant rises in parallel frequency, whereas the frequency of the first formant is lowered. The increases in intensity of the second formant are most pronounced. As the adjustment frequencies of 900 and 1300 Hz address mainly the second formant, a higher impact can be expected.

Furthermore, the intensity of the formant region associated with the singer's formant clusters is increased with increased adjustment frequency (Figure 6). Fant[1] pointed out that small frequency differences between the second formant and the singer's formant cluster may be favorable to higher intensities in the area of 2.5 kHz.

As small formant bandwidths are useful for better intelligibility or articulation,[23] the tuning of the vocal tract toward an external sound might help to enhance vocal capabilities in spoken language as well. Furthermore, professional singers use resonance tuning strategies to increase the audibility of their voices.[24] This mechanism is used mainly at higher f_0 range by tenors and sopranos. In the case presented, a comfortable pitch region with comparably narrow spacing of the harmonics was chosen. From these findings, it seems possible to adjust the position of a specific formant to achieve a desired amplification by the proposed biofeedback.

CONCLUSION AND OUTLOOK

In this study, the first findings on the influence of an adjustment of the vocal tract resonance conditions toward external acoustical signals via biofeedback are presented. The results are in agreement with the model of a driven resonator where the driver is represented by the external sound source and the resonator by the vocal tract. From measurements of phase and amplitude of the enhanced amplitudes, a narrowing of the formant bandwidths after adjustment is deduced. The increased amplitude of the resonance may be linked to lower damping at the frequencies considered by changes of the area function of the vocal tract. The low acoustical resistance of the vocal tract obtained via adjustment of the vocal tract shape without using the vocal folds should therefore facilitate voice production under phonation as well. Furthermore, the adjusted vocal tract preserves its acoustic features during exhalation and even phonation. Thus, highly specific amplification of single frequencies by an adjustment of the vocal tract shape may be helpful in vocal training or warm-up procedures. Moreover, the proposed adjustment may be helpful to reduce the complexity in vocal training and increase the understanding of the basic elements of voice production. Further studies need to be performed to define and evaluate procedures possibly suitable for voice development.

Acknowledgment

The authors thank Thomas Waschkies from Fraunhofer IZFP for helpful discussion. The authors are grateful for the support in experimental instrumentation by Peter Brandtstaedt and Zlatko Dubovski from Fraunhofer-IBP in Stuttgart/Germany and of Karsten Buse from the Fraunhofer-IPM in Freiburg/Germany. Thanks are due to Matthias Struck for access to the audio facilities at the Fraunhofer-IIS.

REFERENCES

1. Fant G. *Acoustic Theory of Speech Production*. Hague, The Netherlands: Mouton; 1970.
2. Lindblom B, Sundberg J. The human voice in speech and singing. In: Rossing TD, ed. *Springer Handbook of Acoustics*. New York: Springer-Verlag; 2007:669–712.
3. Gilman M, Johns MM. The effect of head position and/or stance on the self-perception of phonatory effort. *J Voice*. 2017;31:131.e1–131.e4.
4. Döllinger M, Kaltenbacher M. Recent advances in understanding the human phonatory process. *Acta Acust*. 2016;102:195–208.
5. Mendes AP, Rothman HB, Sapienza C, et al. Effects of vocal training on the acoustic parameters of the singing voice. *J Voice*. 2003;17:529–543.
6. Kob M, Henrich N, Herzel H, et al. Analysing and understanding the singing voice: recent progress and open questions. *Curr Bioinform*. 2011;6:362–374.
7. Maryn Y, De Bodt M, Van Cauwenberge P. Effects of biofeedback in phonatory disorders and phonatory performance: a systematic literature review. *Appl Psychophysiol Biofeedback*. 2006;31:65–83.
8. Lidblom DEF, Sundberg JEF. Acoustical consequences of lip, tongue, jaw, and larynx movement. *J Acoust Soc Am*. 1971;50:1166–1179.
9. Fujimura O, Lindquist J. Sweep tone measurements of vocal tract characteristics. *J Acoust Soc Am*. 1971;49:541.
10. Takemoto H, Mokhtari P. Acoustic analysis of the vocal tract during vowel production by finite-difference time-domain method. *J Acoust Soc Am*. 2010;128:3724–3738.
11. Rothenberg M. Acoustic reinforcement of vocal fold vibratory behavior in singing. In: Fujimura O, ed. *Vocal Physiology: Voice Production, Mechanisms and Functions*. New York: Raven Press; 1988:379–389.
12. Dickens P, Smith J, Wolfe J. Improved precision in measurements of acoustic impedance spectra using resonance-free calibration loads and controlled error distribution. *J Acoust Soc Am*. 2007;121:1471–1481.
13. Epps J, Smith JR, Wolfe J. A novel instrument to measure acoustic resonances of the vocal tract during speech. *Meas Sci Technol*. 1997;8:1112–1121.
14. Stevens NK. *Acoustic Phonetics*. Cambridge, MA, London: The MIT Press; 1998.
15. Wolfe J, Garnier M, Smith J. Vocal tract resonances in speech, singing, and playing musical instruments. *HPFS J*. 2009;3:6–23.
16. Kallail KJ, Emanuel FW. An acoustic comparison of isolated whispered and phonated vowel samples produced by adult male subjects. *J Phonetics*. 1984;12:175–186.
17. Sundberg J. Articulatory interpretation of the singing formants. *J Acoust Soc Am*. 1974;55:838–844.
18. Hanna N, Smith J, Wolfe J. Frequencies, bandwidths and magnitudes of vocal tract and surrounding tissue resonances, measured through the lips during phonation. *J Acoust Soc Am*. 2016;139:2924.
19. Titze IR. Voice training and therapy with a semi-occluded vocal tract: rationale and scientific underpinnings. *J Speech Lang Hear Res*. 2006;49:448–459.
20. Fric M, Hruska V. The effect of resonance tubes on facial and laryngeal vibration—a case study. *Biomed Signal Process Control*. 2017;37:50–60.
21. Denardo B. Design and measurements of variably nonuniform acoustic resonators. *Am J Phys*. 1996;64:745–751.
22. Swerdlin Y, Smith J, Wolfe J. The effect of whisper and creak vocal mechanisms on vocal tract resonances. *J Acoust Soc Am*. 2010;127:2590–2598.
23. Summerfield Q, Foster J, Tyler R, et al. Influences of formant bandwidth and auditory frequency selectivity on identification of place of articulation in stop consonants. *Speech Commun*. 1985;4:213–229.
24. Joliveau E, Smith J, Wolfe J. Tuning of vocal tract resonance by sopranos. *Nature*. 2004;427:116.