One Singer, Two Voices

Johan Sundberg, Björn Lindblom, and Anna-Maria Hefele

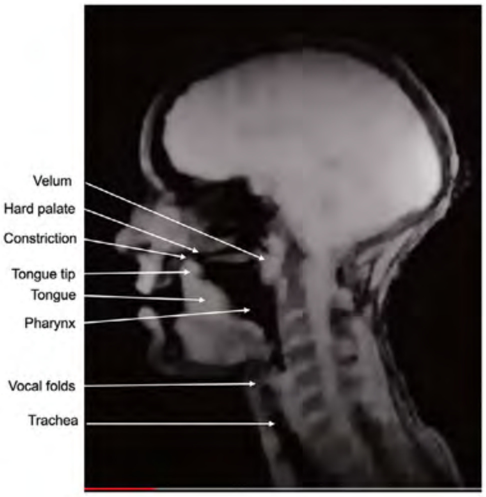
Introduction: Rendering Melodies with Overtones

A single singer but two voices? Experience that situation by visiting world-voice-day.org/EDU/Movies and check the second movie with the title "*Sehnsucht nach dem Frühlinge* (Mozart) — Anna-Maria Hefele (AMH). There, coauthor AMH sings a song by Mozart, first with her singing voice and then with two simultaneous voices, a drone (a low-pitched, continuously sounding tone) plus a whistle-like high-pitched tone that renders the melody. How is this possible? That is the question that we pose here. Let us start by recalling how sounds are created by the instrument AMH is playing, the human voice.

Vocal Sound Generation

Figure 1 shows a frame from the movie mentioned above. It shows a magnetic resonance imaging (MRI) with the various parts of the voice organ labeled. Voice production is the summed result of three processes: (1) compression of air below the vocal folds; (2) vocal fold vibration, quasi-periodically chopping airflow from the subglottal region; and (3) filtering of the acoustic signal of this pulsatile airflow.

Figure 1. *Magnetic resonance (MR) image of Anna-Maria Hefele's (AMH's) head and throat, taken from the video where she performs a Mozart melody in overtone singing technique.*



The overpressure of air below the folds throws them apart, thus allowing air to pass through the slit between them. Then, aerodynamic conditions reduce the air pressure along the folds, which, together with the elasticity of their tissue, closes the slit. The same pattern is then repeated, thus generating vocal fold vibration.

The vibration generates a pulsatile airflow as seen in **Figure 2A**, producing sound, the voice source. The pitch is determined by the vibration frequency, whereas the waveform is far from sinusoidal. Hence, this airflow signal is composed of a number of harmonic partials. In other words, the frequency of a partial number (n) = $n \times f_o$, where f_o is the frequency of the lowest partial, the fundamental or vibration frequency. The amplitudes of the partials tend to decrease with their frequency; the amplitude of n tends to be something like 12 dB stronger than the amplitude of $n \times 2$. The spectrum envelope of the voice source is rather smooth and has a negative slope as seen in **Figure 2B**.

The voice source is injected into the vocal tract (VT), which is a resonator. Hence it possesses resonances at certain frequencies. Partials with frequencies close to a VT resonance frequency are enhanced and partials further away are attenuated (see **Figure 2C**). Therefore, the spectrum envelope of the sound radiated from the lip opening (**Figure 2A**) contains peaks at the VT resonance frequencies and valleys in-between them. In this sense, the VT resonances form the spectrum envelope of the sound emitted to the free air. Probably for this reason, VT resonances are frequently referred to as formants.

The frequencies of the formants are determined by the shape of the resonator composed of the pharynx and the mouth cavities, the VT. For example, the VT length has

https://doi.org/10.1121/AT.2021.17.1.43

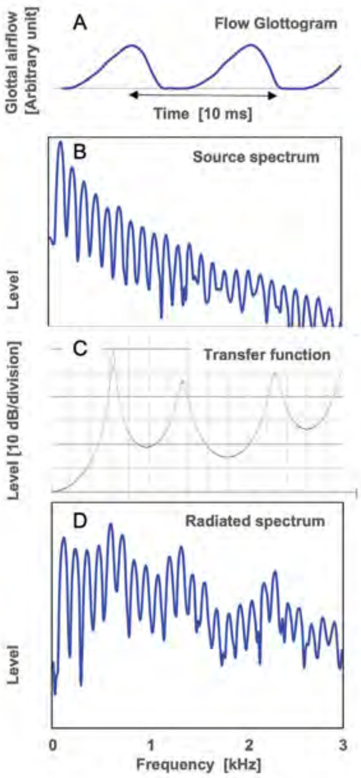


Figure 2. *A and B: typical waveform and spectrum, respectively, of the glottal airflow during phonation.* ***C*** *and* ***D****: vocal tract transfer function for the vowel /ae/ and the corresponding radiated spectrum, respectively.*

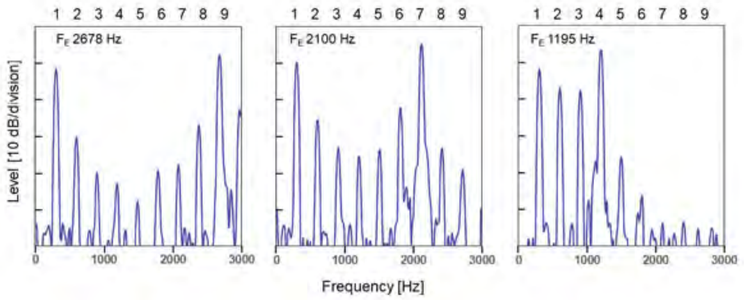
a strong effect on these frequencies and protruding the lips makes the VT longer, thus lowering the formant frequencies. The shape of the VT can be varied within wide limits. Moreover, by bulging the tongue more or less and in various directions, the VT can be narrowed or widened in almost any place along its length axis from the deep pharynx to the hard palate. Also, the jaw and lip openings contribute to determining the VT shape. As a result, the formant frequencies can be varied within quite wide ranges: the first formant between about 150 and 1,000 Hz, the second from about 500 and 3,000 Hz, and the third from about 1,500 Hz and 4,000 Hz.

Overtone Singing

What is overtone singing, then? The term covers several different styles. Overtone singing was described as early as the nineteenth century by a famous singing teacher Manuel Garcia (Wendler, 1991) and has attracted the interest of several researchers. Smith and colleagues (1967) described "throat singing," a type of chant performed by Tibetan Lamas, (see, e.g., s.si.edu/37QCSOZ). It is produced by males with special types of vocal fold vibrations, referred to as vocal fry register. Its pitch is stable and very low and is produced by a vocal fold vibration pattern in which every second or third airflow pulse is attenuated. Consequently, the pitch period of this drone is doubled or tripled. A similar type of phonation often occurs in phrase endings in conversational speech but is then typically aperiodic. In throat singing, two of the overtones are quite strong, and audible, thereby together giving the impression of a "chord." Throat singing is regarded as sacred in some Asian cultures.

The overtone singing demonstrated by AMH in the above link can be produced by both females and males. However, the fundamental frequency of the drone is not as low as in throat singing. In AMH's case, it is in the range typical of female speaking voices. The melody is played in a much higher pitch range by very strong overtones. **Figure 3** shows some examples where overtones number 9, 7, and 4 are the strongest in the spectrum.

Figure 3. *Examples of spectra produced in overtone singing by AMH. F*E*, frequency of the enhanced overtone.*



In everyday listening, overtones are not perceived individually. Instead, the patterning of their amplitudes, determined by the resonance characteristics of the VT, collectively contributes to what our auditory system perceives as the "timbre" or vocal "color" of what is sung or spoken. The reason why overtones normally escape us is linked to the way our hearing works. It processes spectral contents by averaging the information in broad frequency bands, the so-called critical bands (Moore, 2012).

Characteristics of Overtone Singing

Against this background, we may be forgiven for finding overtone singing a rather puzzling phenomenon. Here we present an attempt to shed some light on its phonatory, articulatory, and acoustic bases.

We begin with a sample from AMH's rendering of the Mozart melody. **Figure 4A** presents the first few bars of the beginning of Mozart's theme in musical notation. Let us take a moment to consider what it would take to sing this sequence using overtones?

Figure 4 gives an answer in-principle. The format of the musical score is used to indicate the timing and pitch of each overtone. Along the frequency scale, the first eight harmonics are drawn at equidistant intervals. Hypothetically, let us suppose that the singer selects a fundamental frequency near 300 Hz for the drone. That implies a "keyboard" of the following frequencies of the first eight overtones: #2, 600 Hz; #3, 900 Hz; #4,1,200 Hz; #5,1,500; #6, 1,800 Hz; #7, 2,100 Hz; and #8, 2,400 Hz.

Note the following ratios: (1) 1,500/1,200 = 1.25; (2) 1,800/1,200 = 1.5; (3) 2,400/1,200 = 2. Relative to the first note at 1,200 Hz, the ratios correspond to a major third, a perfect fifth, and an octave, respectively. Those intervals are the ones needed to produce the notes of the first two bars.

For readers who find that result just a little too convenient, we should point out that it is no accident. Our musical scales and their intervals bear a very close evolutionary relationship to the physical structure of periodic sounds. Such sounds are constituted by the partials, the frequencies of which form a harmonic series (Gill and Purves, 2009). This implies that the musical intervals octave, fifth, fourth, major third, and minor third appear between the six lowest spectrum partials. Hence, it is possible to play melodies with the partials of a constant drone tone.

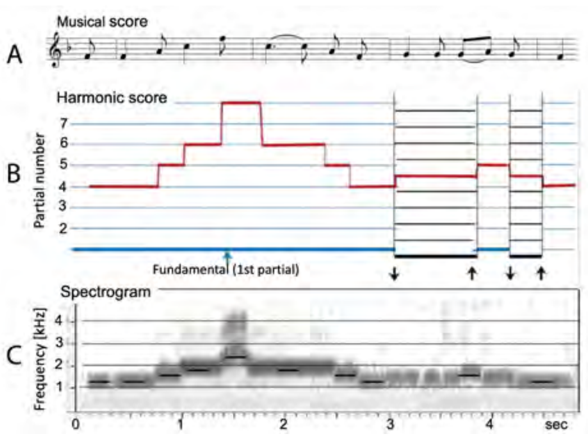


Figure 4. *Schematic illustration of singer AMH's overtone singing performance of a Mozart song.* ***A:*** *the musical score.* ***B:*** *partials used for the melody tones (****red lines****) and for the drone (****blue*** *and* ***black lines****).* ***C:*** *wideband spectrogram of the performance.*

However, in the third bar in **Figure 4A** with chord tones C E G Bb, the pitch G appears. The series just mentioned does not provide an equally convenient choice for that note. Looking ahead to AMH's performance (**Figure 4B**), we see how she handles the situation: She lowers the fundamental of the drone so as to produce an overtone whose relationship to 1,200 Hz is that of a major second or about 1,125 Hz! Our harmonic score includes that approach as illustrated by a shift in the fundamental and harmonics of the drone.

Figure 4C shows real data, a spectrogram of one of AMH's overtone singing versions of the song. To enhance the display of the overtones, we show a wideband filtering that portrays the overtones as dark patches. To help interpret their positions, we added black short lines that indicate the expected frequency values on the assumption that the 4th, 5th, 6th, and 8th harmonics of a 300-Hz fundamental serve as the melodic building blocks.

We note that these predictions parallel AMH's overtones rather well. However, the black marks slightly underestimate the observed values. Why? AMH used a fundamental frequency slightly lower than our hypothetical 300 Hz.

This example illustrates the fact that overtone singing derives from the lawful way in which the harmonics are organized in periodic sounds. Overtone singers are able to exploit this patterning. They have developed a way of selecting and amplifying the harmonics and have refined their VT motor skills to be able, with great precision in time and frequency, to produce harmonics as melodic sequences. Next, we test the hypotheses that (1) enhancing and selecting a single partial is the result of VT shapes producing clustering of formants; and (2) that overtone singing is produced with a regular sound source.

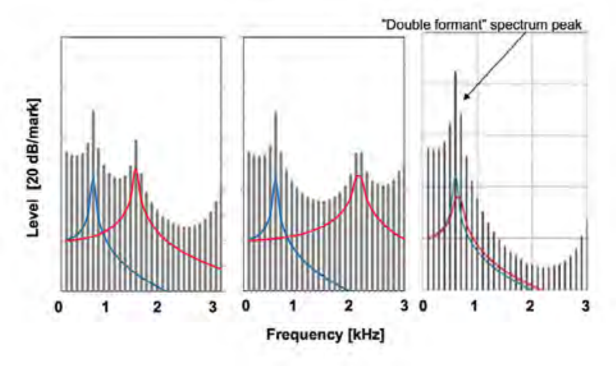
Acoustic Theory

Enhancing Single Harmonics

How can formants produce the excessive amplitudes of the single overtones illustrated in **Figure 3**? Mathematically, the spectral envelope, the function determining the amplitudes of the overtones, is the sum of formant resonance curves (which vary with changes in VT shape) and certain factors such as glottal source and radiation characteristics (which do not depend on articulatory activity). The only input to the calculation is the frequencies and bandwidths of the formants. The latter factor generally varies with the former factor in a predictable manner so formant amplitudes need not be specified. **Figure 5** illustrates this predictability.

Figure 5 shows three line spectra with the cardinal shapes of the resonance curves for the first and second formants (henceforth F1 and F2). The amplitudes of the partials and their spectral envelopes were derived in accordance with the standard source-filter model (Fant, 1960). This theory treats the envelope as the sum of formant curves and the constant contributions of source and radiation characteristics, which are not shown in **Figure 5**. The bandwidths are secondary aspects, being determined mainly by the frequencies of the formants (Fant, 1972).

Figure 5. *Schematic illustration of the spectrum effects of moving the frequencies of two formants closer together.* ***Vertical lines****, partials of a drone with a fundamental frequency of 100 Hz;* ***blue*** *and* ***red curves****, first (F1) and second (F2) formants, respectively.* ***Left:*** *F1 = 600 Hz, F2 = 1,400 Hz.* ***Center:*** *F1 = 600 Hz; F2 = 2,150 Hz.* ***Right:*** *F1 = 600 Hz; F2 = 650 Hz, thus creating a "double formant," that creates a very strong partial (****arrow****).*



In **Figure 5**, F1 was fixed at 600 Hz while F2 was varied. When F1 and F2 approach almost identical frequencies (**Figure 5**, *right*), creating, as it were, a *double formant*, their individual peaks merge into a single maximum, with a significant increase in the amplitude of the closest partial. In other words, acoustic theory states that formant amplitudes are predictable and thus suggests an answer to the question asked in the first sentence of this section. Enhancing the amplitude of individual overtones is possible: *Move two formants close to each other in frequency. Create a double formant!*

Measurements and Modeling

Vocal Tract Shapes

As shown above, if the formant frequencies are determined by the shape of the VT, so what was the shape of AMH's VT? This has actually been documented in another dynamic (MRI) video published by the Freiburg Institute of Musician's Medicine, *Naturtonreihe in Zungentechnik* (see youtu.be/-jKl61Xxkh0). It was taken when AMH performed overtone singing, enhancing, one by one each overtone of a drone with a fundamental frequency of 270 Hz (pitch about C4), in a rising followed by a descending sequence. Henceforth, the frequencies of the enhanced overtones will be referred to as F_E. All overtones, from the 4th, F_E = 1,080 Hz, up to the 12th, $F_E \approx$ 3,200 Hz, were enhanced.

The MRI video shows her entire VT in a midsagittal lateral profile. **Figure 6** shows tracings of the VT for each of the enhanced overtones in the ascending and the descending series.

Voice Source

Is formant clustering an exhaustive explanation of overtone singing? Fortunately, the transfer function of the VT can be predicted given its formant frequencies. Thus, a vowel spectrum can be analyzed not only with respect to the formant frequencies, which appear as peaks in the spectrum envelope, but also with respect to the voice source. The trick is simple, inverse filtering!

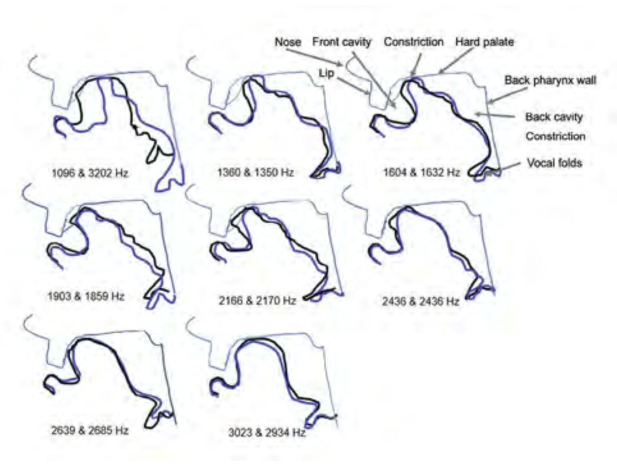


Figure 6. *Tracings of the lateral midsagittal articulatory profiles observed in the MR images of the subject while producing the indicated values of* F_E *in the ascending (**blue**) and descending (**black**) sequences.*

On its way from the glottis to the lip opening, the voice source has been filtered by the transfer function of the VT. Inverse filtering means that the radiated spectrum is filtered by the VT transfer function (**Figure 2C**) turned upside down.

The transfer function itself can be computed from the formant frequencies. This may sound a bit circular, but it is not. Glottal airflow must be zero when the glottis is closed. Hence the closed phase of the voice source waveform contains no ringing at a formant frequency if the inverse filter exactly equals the transfer function. Moreover, it is well-known that the spectrum envelope of the voice source has a smooth spectrum envelope, so peaks and valleys close to the formant frequencies are signs of inaccurate tuning of the inverse filters. Thus fine tuning of the inverse filters is a condition for reaching an accurate result. Errors reveal themselves in terms of ringing during the closed phase and/or a spectrum envelope peak and/or a trough near the formants.

The voice source can be varied along three dimensions. By stretching and tensing the vocal folds, the fundamental frequency increases, resulting in an increase in pitch. By increasing the overpressure of air in the respiratory system, the amplitude of the voice source increases, which causes vocal loudness to increase. By changing vocal fold adduction, which results in squeezing the glottis, the voice timbre varies along a dimension that ranges from breathy to pressed. Breathy phonation is what you typically use during a concert when you want to tell something to the person sitting next to you without disturbing the performance. Pressed phonation is typically used when you speak in excited anger or when you attempt to say something when carrying something very heavy.

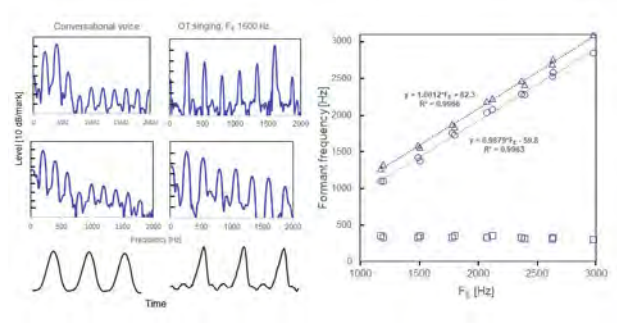
Figure 7 shows two examples of AMH's voice. Compared with her conversational speech, the source spectrum envelope slopes less steeply in overtone singing. Furthermore, the waveform has a longer closed phase and contains sharp knees, typical signs of an increase of vocal fold adduction. The ripple during the closed phase on overtone singing does not correspond to a formant frequency, but to a 900-Hz periodicity, an artifact frequently observed in glottal flow waveforms.

The frequencies of the three lowest formants used for the inverse filtering are plotted as function of F_E in **Figure 7**, *right*. The trend lines show that F2 and F3 have similar slopes and intercepts differing by about 185 Hz. Thus F2 and F3 are closely clustered around F_E, suggesting an affirmative answer to the question raised above, if formant clustering is the sole explanation of overtone singing. As formant frequencies are controlled by the shape of the VT, the next question then is how AMH shapes her VT to achieve this distribution of formant frequencies.

Estimating Vocal Tract Shapes

The resonances of the VT are determined by its shape, and we have excellent tools for varying this shape within

Figure 7. *Examples of AMH's voice in conversational speech (**left**) and during overtone (OT) singing (**center**). **Top:** radiated spectra. **Center:** voice source spectra; **Bottom:** glottal airflow waveforms. **Right:** the three lowest formant frequencies used for the inverse filtering of AMH's overtone singing as functions of* F_E*. **Lines** and **equations** represent trend lines.*



very wide limits. We can vary the shape of the tongue body, the position of the tongue tip, the jaw and lip openings, the larynx height, and the position and status of the gateway to the nose, the velum.

Let us now more closely examine the shape of AMH's VT as documented in the MRI video. It is evident from **Figure 6** that AMH produced overtone singing with a lifted tongue tip, so the tongue tip divided the VT into a front cavity and a back cavity. Our first target is the back cavity posterior to the raised tongue tip.

The formant frequencies associated with a given VT shape can be estimated from the VT contour. Several investigations have examined the relationship between the sagittal distance separating the VT contours and the associated cross-sectional area at the various positions along the VT length axis (see, e.g., Ericsdotter, 2005). Hence it was possible to describe the shape of the back cavity for each FE in terms of an area function that lists the cross-sectional area as a function of the distance to the vocal folds.

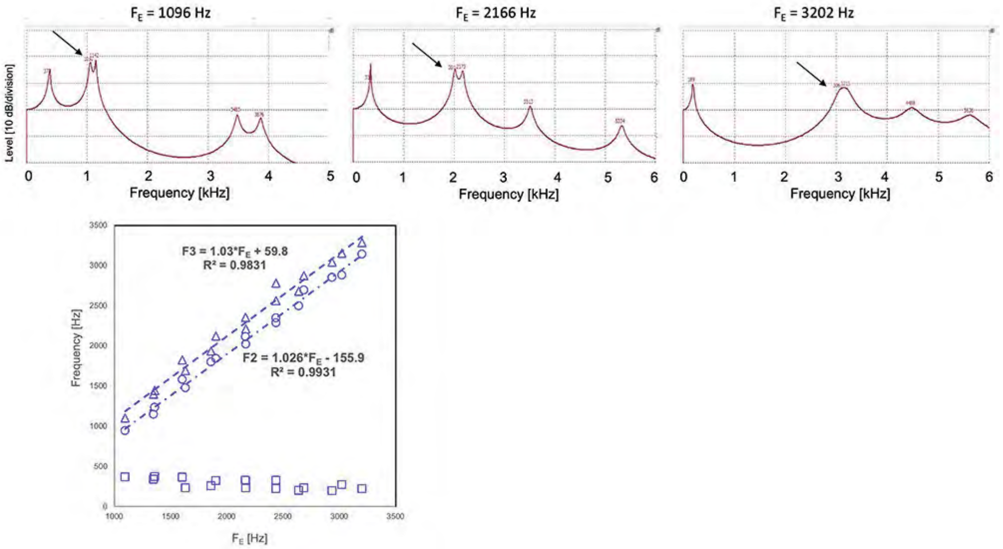
The next question concerns the front cavity, anterior to the raised tongue tip. The cavity between palatal constriction and the lip opening looks like, and can be regarded as, a Helmholtz resonator, a cavity in front of the tongue tip and a neck formed by the rather narrow lip opening (Sundberg and Lindblom, 1990; Granqvist et al., 2003).

The area of the lip opening was measured in a front video recorded when AMH produced the same overtone series as for the MRI video. The length of the lip opening was documented in the MRI video. These measures plus the frequency of the third formant used for the inverse filtering analysis allowed us to use the Helmholtz equation for calculating the front cavity volume. The validity of this approximation was corroborated in terms of a strong correlation between the measured length and the volume of the front cavity.

The formant frequencies of the entire VT could be calculated by a custom-made software, *Wormfrek* (Liljencrants and Fant, 1975). **Figure 8** shows the transfer functions with the formant frequencies for three FE values: 1,096, 2,166, and 3,202 Hz. In **Figure 8**, the *arrows* highlight the close proximity of F2 and F3. In **Figure 8**, *bottom*, F1, F2, and F3 are plotted as a function of FE. The trend lines show that F2 and F3 have similar slopes and intercepts differing by about 220 Hz.

We note that the F1, F2, and F3 predictions parallel the formant measurements made using inverse filtering (**Figure 7**). Here, a somewhat wider distance separates F2 from F3 than what was shown in **Figure 7**. The common

Figure 8. ***Top:*** *Wormfrek software displays of the transfer functions for the lowest, a middle, and the highest FE (**left, center,** and **right,** respectively).* ***Bottom:*** *associated values of F1, F2, and F3 as a function of FE.* ***Lines*** *and* ***equations*** *refer to trend line approximations.*



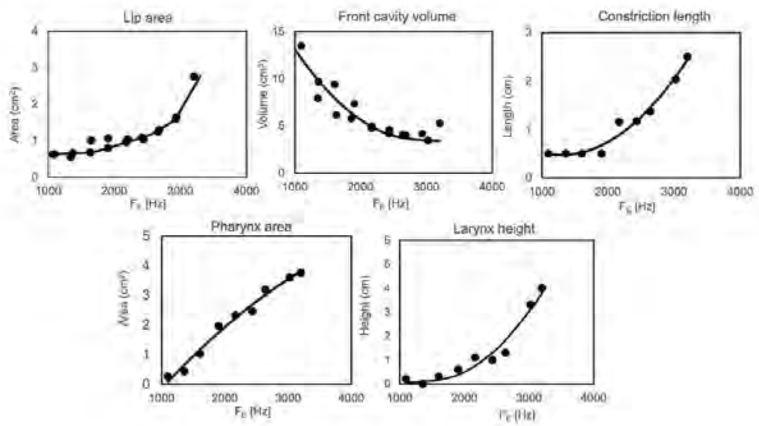


Figure 9. *Articulatory area function parameters plotted as functions of* FE. **Curves** *show approximations derived from trend line equations.*

denominator is the consistent identification of the double formant. We feel justified in concluding that our results confirm the double formant phenomenon as a prerequisite for the overtone selection and enhancement in AMH's overtone singing technique.

Conclusions

Central to the present account is the "double formant" hypothesis, which attributes the phenomenon of overtone singing to VT filtering. However, the inverse filtering results also suggest that overtone singing involves a phonation type different from that in conversational voice, making the source spectrum slope less steep and thus boosting the amplitudes of the higher overtones. These findings replicate and extend previous investigations of overtone singing. Bloothooft et al. (1992) undertook an acoustic study of an experienced overtone singer and suggested formant clustering as an explanation and also noted an extended closed phase of the vocal fold vibrations. Using impedance measurements, Kob (2004) analyzed a form of overtone singing called sygyt and interpreted the overtone boosting as the result of formant clustering.

Parallel vibrations of the ventricular folds have been documented in throat singing (Lindestad et al., 2001). How about this possibility in AMH's overtone singing? Our inverse filtering data clearly rule out the existence of a laryngeal mechanism that selectively amplifies and enhances individual partials.

Overtone singing clearly requires an extremely high degree of articulatory precision; for each FE, two cavities need to be shaped such that they produce resonance frequencies that match each other within a few tens of Hertz. How can the underlying motor control be organized? It is probably relevant that some of the articulatory configurations shown in **Figure 6** are used also in speech. The lateral profile for FE = 1,096 Hz resembles the articulation of retroflex consonants (Dixit 1990; Krull and Lindblom, 1996). A narrow pharyngeal constriction is typical of [a]-like vowels and pharyngealized consonants (Ladefoged and Maddieson, 1996). The VT for FE = 3,202 Hz has a "palatalized' tongue shape similar to that used for the vowel [i].

It would also be relevant that the articulatory parameters varied systematically with FE. This is illustrated in **Figure 9**. It shows how AMH varied the lip opening area, length of palatal constriction, larynx height, front cavity volume, and pharynx area as a function of FE. It is evident that the values of each individual articulatory dimension are aligned along smooth contours running between its values in FE = 1,096 and 3,202 Hz. This lawful patterning suggests that it would be possible to derive VT shapes intermediate between those for FE = 1,096 and 3,202 Hz by interpolation. A rough description would be to say that the VT shapes are located along a trajectory in the articulatory space that runs between a retroflex and pharyngealized [a] and an [i]-like, palatalized tongue profile.

How to Learn Overtone Singing

Producing overtones with your own voice is relatively easy. You practice singing very slow vowel transitions between the vowels /i/ and /u/ on a long-sustained drone that is kept at a constant pitch. Then, overtones start to appear quite clearly from your voice, although you might not able to hear them yet. To hear overtones in your own voice is the key to achieving deliberate control; learning to hear them is the first important part of your practice.

In this article, we have analyzed an advanced technique of overtone singing, double resonator articulation. The tongue tip is retracted and elevated in the mouth as for the American consonant /r/. This lowers the third formant and can bring it close to the second formant. As we have seen, this creates a double resonator and a double formant, which results in a strong, whistling-like overtone. To do this requires quite an accurate and simultaneous control over the front cavity for the third formant and the back cavity for the second formant. Generally, it takes quite some practice to learn this technique.

A simpler start into the fascinating world of overtone singing may be to learn to enhance overtones with vowels only, with an undivided VT cavity. Then, the VT works as a single resonator, and the second formant is solely responsible for overtone enhancement. Also, this technique can be learned by very slowly changing the articulation between /i/ and /u/, keeping a drone with constant pitch. When you manage to do this, you will discern single overtones; one by one, they first increase and then decrease in loudness as they approach the second formant, pass it, and then move away from it, and soon after, the next overtone will appear and do the same thing.

After you have learned the vowel-technique well, it is mostly both exciting and not too difficult to learn the double formant technique. Then, you may want to explore the pleasure of shifting the drone pitch and so extend the melodic possibilities of overtone singing even further. If you want to learn more, see Hefele (2020)!

Acknowledgments

Bernhard Richter, Matthias Echternach, Louisa Traser, and Michael Burdumy at the Freiburg Institute of Musician's Medicine are acknowledged for their work with the two dynamic magnetic resonance imaging videos.

References

Bloothooft, G., Bringmann, E., van Cappellen, M., van Luipen, J. B., and Thomassen, K. P. (1992). Acoustics and perception of overtone singing. *The Journal of the Acoustical Society of America* 92, 1827-1836.

Dixit, R. P. (1990). Lingotectal contact patterns in the dental and retroflex stops of Hindi, *Journal of Phonetics* 11, 291-302.

Ericsdotter, C. (2005). *Articulatory-Acoustic Relationships in Swedish Vowel Sounds*, Doctoral Dissertation, Stockholm University, Stockholm, Sweden.

Fant, G. (1960). *The Acoustic Theory of Speech Production*, Mouton, The Hague, The Netherlands.

Fant, G. (1972). Vocal tract wall effects, losses and resonance bandwidths. *Quarterly Progress and Status Report* 2(3), Department of Speech Hearing and Music (KTH), Stockholm, Sweden, pp. 28-52. Available at https://rb.gy/ugtss4.

Gill, K. Z., and Purves, D. (2009). A biological rationale for musical scales. *PLoS ONE* 4(12): e8144.

Granqvist, S., Sundberg, J., Cortes, E. E., Larsson, J., and Branderud, P. (2003). The front and sub-lingual cavities in coronal stops: An acoustic approach to volume estimation. *Proceedings of the 15th International Congress of Phonetic Sciences*, Barcelona, Spain, August 3-9, 2003, pp. 941-944.

Hefele, A.-M. (2020). Anna-Maria Hefele's Website www.overtone.academy, Available at: https://rb.gy/yplwko. Accessed: October 12, 2020.

Kob, M. (2004). Analysis and modelling of overtone singing in the sygyt style. *Applied Acoustics* 65 1249-1259.

Krull, D., and Lindblom, B. (1996). Coarticulation in apical consonants: acoustic and articulatory analyses of Hindi, Swedish and Tamil, *Quarterly Progress and Status Report 2*, Department of Speech Hearing and Music (KTH), Stockholm, Sweden, pp. 73-76. Available at https://rb.gy/ugtss4.

Ladefoged, P., and Maddieson, I. (1986). *The Sounds of the World's Languages*, Blackwell, Oxford, UK.

Liljencrants, J., and Fant, G. (1975). Computer program for VT-resonance frequency calculations, *Quarterly Progress and Status Report 4*, Department of Speech Hearing and Music (KTH), Stockholm, Sweden, pp. 15-20. Available at https://rb.gy/ugtss4.

Lindestad, P.-Å., Södersten, M., Merker, B., and Granqvist, S. (2001). Voice source characteristics in Mongolian "throat singing" studied with high-speed imaging technique, acoustic spectra, and inverse filtering, *Journal of Voice* 15, 78-85.

Moore, B. C. J. (2012). *An Introduction to the Psychology of Hearing*, Emerald Group Publishing Ltd,. Bingley, UK

Smith, H., Stevens, K., and Tomlinson, R. (1967). On an unusual mode of singing by certain Tibetan lamas. *The Journal of the Acoustical Society of America* 41, 1262-1264.

Sundberg, J., and Lindblom, B. (1990). Acoustic estimation of the front cavity in apical stops *The Journal of the Acoustical Society of America* 88, 1313-1317.

Wendler J., Dejonckere, P. H., Smit, P. H., and Cebulla, M. (1991). Overtone singing in the view of Manuel Garcia. *Proceedings of Conference "New Ways of the Voice,"* Service ORL et Audiophonologie, Besancon, France, February 14-17, 1991, p. 46.

About the Authors

Johan Sundberg jsu@kth.se

Department of Speech Music
and Hearing
Royal Institute of Technology (KTH)
SE-100 44 Stockholm, Sweden

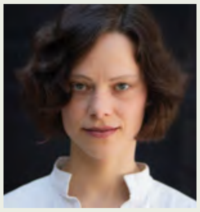
Johan Sundberg studied musicology at Uppsala University (Uppsala, Sweden). He made his dissertation research on the acoustics of organ pipes as a guest researcher in Gunnar Fant's department at the Royal Institute of Technology (KTH; Stockholm, Sweden). This brought him into a productive contact with Björn Lindblom. After finishing his dissertation in 1966, he founded a research group in the area of music acoustics at the KTH and was awarded a personal chair in music acoustics there in 1979. Being an active singer, the voice as a musical instrument has been his main research theme along with the theory underlying music performance.

Author photo by Linnéa Heinerborg

Björn Lindblom
lindblom@ling.su.se

Department of Linguistics
Stockholm University
SE-106 91 Stockholm, Sweden

Björn Lindblom became an experimental phonetician in the early 1960s. His publications span a wide range of topics, including the development, production, and perception of speech. Academic experience: teaching and doing laboratory research at the Royal Institute of Technology (KTH; Stockholm, Sweden), Haskins Laboratories (New Haven, CT), MIT (Cambridge, MA), Stockholm University (SU), and the University of Texas at Austin (UT). He has held endowed chairs at SU and UT. He is a Fellow of the Acoustical Society of America and of the American Association for the Advancement of Science (AAAS) and is an Honorary Life Member of the Linguistic Society of America (LSA). His current project is a book: *Reinventing Spoken Language — The Biological Way*.



Anna-Maria Hefele
ama@anna-maria-hefele.com

Overtone Academy
Saulengrainerstrasse 1
DE-87742 Dirlewang, Germany

Anna-Maria Hefele has a Master of Arts from Mozarteum Salzburg (Austria) and is a multi-instrumentalist singer and overtone singer, performing worldwide as a soloist with different ensembles, choirs, and orchestras. She frequently performs in contemporary ballet, circus, and dance theater productions. Her YouTube video "Polyphonic Overtone Singing" went viral and has resulted in more than 17 million views so far, followed by regular appearances in various international television shows and radio broadcasts. Headlines like "A Voice as from Another World," "The Lady with the Two Voices," and "Polyphonic Vocalist Does the Impossible" have spread across the world.

Author photo by Thomas Radlwimmer