Freddie Mercury—acoustic analysis of speaking fundamental frequency, vibrato, and subharmonics

Christian T. Herbst, Stellan Hertegard, Daniel Zangger-Borch & Per-Åke Lindestad

To cite this article: Christian T. Herbst, Stellan Hertegard, Daniel Zangger-Borch & Per-Åke Lindestad (2016): Freddie Mercury—acoustic analysis of speaking fundamental frequency, vibrato, and subharmonics, Logopedics Phoniatrics Vocology, DOI: 10.3109/14015439.2016.1156737

To link to this article: http://dx.doi.org/10.3109/14015439.2016.1156737

Published online: 15 Apr 2016.
**Freddie Mercury—acoustic analysis of speaking fundamental frequency, vibrato, and subharmonics**

Christian T. Herbst*, Stellan Hertegardb, Daniel Zangger-Borchc and Per-Åke Lindestadb

*Voice Research Lab, Department of Biophysics, Faculty of Science, Palacký University Olomouc, 17. listopadu 12, 771 46 Olomouc, Czech Republic; bDepartment of Otolaryngology, Head and Neck Surgery B53, Clinical Sciences and Intervention, Karolinska Institutet, Karolinska University Hospital, 141 86 Stockholm, Sweden; cVoice Centre Stockholm AB, Zangger Vocal Art, Box 4342, 102 67 Stockholm, Sweden

**ABSTRACT**

Freddie Mercury was one of the twentieth century’s best-known singers of commercial contemporary music. This study presents an acoustical analysis of his voice production and singing style, based on perceptual and quantitative analysis of publicly available sound recordings. Analysis of six interviews revealed a median speaking fundamental frequency of 117.3 Hz, which is typically found for a baritone voice. Analysis of voice tracks isolated from full band recordings suggested that the singing voice range was 37 semitones within the pitch range of F#2 (about 92.2 Hz) to G5 (about 784 Hz). Evidence for higher phonations up to a fundamental frequency of 1,347 Hz was not deemed reliable. Analysis of 240 sustained notes from 21 a-cappella recordings revealed a surprisingly high mean fundamental frequency modulation rate (vibrato) of 7.0 Hz; reaching the range of vocal tremor. Quantitative analysis utilizing a newly introduced parameter to assess the regularity of vocal vibrato corroborated its perceptually irregular nature, suggesting that vibrato (ir)regularity is a distinctive feature of the singing voice. Imitation of subharmonic phonation samples by a professional rock singer, documented by endoscopic high-speed video at 4,132 frames per second, revealed a 3:1 frequency locked vibratory pattern of vocal folds and ventricular folds.

**Introduction**

Freddie Mercury, born 5 September 1946 as Farrokh Bulsara, and died 24 November 1991, was one of the most influential singers of his time. The *Rolling Stone* magazine listed him eighteenth within the 100 best (commercial contemporary) singers of all time (1). Having been the lead singer of the rock band Queen, he profoundly influenced this group’s musical style over more than two decades.

Freddie Mercury’s voice has been described as ‘a force of nature with the velocity of a hurricane’ (2), which was ‘escalating within a few bars from a deep, throaty rock-growl to tender, vibrant tenor, then on to a high-pitched, perfect coloratura, pure and crystalline in the upper reaches’ (3, p. 2). Such descriptions, while presumably adequate for a biography or a newspaper article, do not satisfy deeper scholarly interest into the singer’s voice characteristics. The purpose of this study was therefore to conduct a viable analysis of publicly available data material, in order to arrive at more empirically based insights into Freddie Mercury’s voice production and singing style.

**Materials and methods**

The task of objectively describing the voice of a singer who is not available for acquiring recordings under controlled acoustical and physiological conditions is ambitious. Special care needs to be taken in selecting proper data material (ideally a cappella, no background sounds) and suitable analysis methods. In this study, a series of exemplary interpretations of voice quality were conducted, and, where deemed appropriate, quantitative estimation of voice features like speaking fundamental frequency or vibrato properties was performed. These analyses, in part inspired by a previous publication by the first author (4), are conducted here with improved methodology and larger data sets. Finally, endoscopic high-speed video examination of vocal fold vibration of a contemporary rock singer mimicking typical sounds produced by Freddie Mercury was performed, in an attempt to understand the voice production mechanism of subharmonic sounds. No formant analysis was undertaken, however, since it cannot be taken for granted that the spectral contents of commercially available recordings have not been considerably changed by the editing sound engineers.

**Speaking fundamental frequency (SFF) and singing voice range**

In order to assess Freddie Mercury’s speaking fundamental frequency (SFF), a total of six interviews, conducted by Andrew Wigg (1984, 1985, 1986, 1987) and Rudi Dolezal...
(1984 and 1987) were analyzed. The interviews by Andres Wigg were taken from the commercial compilation Freddie Mercury: The Solo Collection (2000, EMI/Parlophone). The interview conducted by Andrew Wigg in 1979 was excluded from the analysis due to excessive background noise. The interviews by Rudi Dolezal (‘A musical prostitute’, 1984, and ‘The last interview’, 1987) were downloaded from the internet (https://www.youtube.com/watch?v=8wk9hPubD1Q and https://www.youtube.com/watch?v=PnsiSjsSNYQ).

The acoustic tracks of all interviews were normalized in amplitude and then annotated and analyzed with the software Praat (5). Only those segments were considered where without doubt Freddie Mercury’s voice was exclusively audible. The fundamental frequency of all extracted sequences was estimated using Praat’s autocorrelation algorithm (6), using a rather conservative parameter setting in order to avoid analysis artifacts stemming from background noise or unvoiced consonants: fundamental frequency range: 70–400 Hz; silence threshold: 0.2; voicing threshold: 0.6; octave cost: 0.01; octave jump cost: 0.8; voiced/unvoiced cost: 0.14; read progress: 10 ms. The analysis resulted in a total of 134,418 data points.

For the perceptual assessment of the singing voice range, Freddie Mercury: The Solo Collection as well as 23 commercially available recordings of Queen were considered. The assessment was conducted by author C.T.H. by matching the perceived pitch of candidate sound samples, sung at extremely high and low frequencies, to the perceived musical pitch of candidate sound samples, sung at 440 Hz (equal temperament). Only those sound samples were considered where Freddie Mercury could be clearly identified as the exclusively audible singer, and no background vocals were included in the analysis.

**Descriptive analysis of functional singing voice parameters**

In speech and singing, the spectral composition of the emitted sound is causally determined by the configuration of the sound production organ, constituted by the respiratory system, the larynx, and the vocal tract. At the laryngeal level, two main determinants are constituted by the vocal register and the degree of posterior vocal fold adduction, controlled along the dimension of ‘breathy’ versus ‘pressed’ (7,8).

Without any physiological or biomechanical data of vocal fold vibration the assessment of the laryngeal configuration during singing is difficult, particularly when based on commercial recordings whose spectral composition could have been changed by sound engineers and by influences of the recording equipment (microphone and amplifier transfer functions, data conversion, recording device), room acoustics (reverberation), and the mouth-to-microphone distance (high-frequency selective directionality) (9). Nevertheless, a small number of exemplary sound samples are discussed here, based on perceptual and (partly) spectral analysis, and hypothetical physiological interpretations as regards vocal register and posterior vocal fold adduction are suggested. The perceptual assessments were conducted by the four authors of this manuscript (two graduated voice pedagogues and two phoniatricians), and all assessments are based on unanimous consensus of the four authors.

**Analysis of vocal vibrato**

Vibrato is constituted by a (predominantly regular) modulation of the singing fundamental frequency, typically in the range of 4–7 Hz (10). For the assessment of Freddie Mercury’s vibrato characteristics, the compilation The Acappella Collection (http://queenspain.blogspot.co.at/2013/06/freddie-mercury-acapella-collection.html), compiled from multichannel recordings by the Queen Poland fan site (https://queenpoland.wordpress.com/), was analyzed. This assortment of studio recordings contains the vocal tracks of 28 songs by Queen and Freddie Mercury. Sustained notes from these recordings have been considered for analysis when the following conditions were met: 1) exactly one singing voice was audible, and this voice could be clearly identified as Freddie Mercury’s; 2) no background noise, background vocals, or orchestration was audible; 3) no noteworthy delay or reverberation effects have been applied; and 4) the sustained note had a duration of more than 500 ms and contained audible cyclic frequency modulations.

Data annotation and f0 extraction was conducted with the software Praat. For f0 extraction, Praat’s autocorrelation algorithm was used with these parameters: timeStep: 10 ms; frequency range 70–600 Hz; voicingThreshold: 0.3; octaveJumpCost: 0.35; silenceThreshold: 0.03; octaveCost: 0.07; voicedUnvoicedCost: 0.14. After the automated analysis, all extracted fundamental frequency contours were superimposed upon the respective spectrograms and then visually assessed for correct f0 extraction. Nine segments were excluded from the analysis since they contained obvious errors (mostly octave jump errors during the occurrence of subharmonics).

A total of 128 segments of sustained notes (from 15 of the 28 songs in The Acappella Collection) with vibrato (11,475 data points total, 114.75 seconds total duration), having reliably estimated f0 contours of over 500 ms duration each, were considered for further analysis. For each segment, three quantities were computed: 1) the average fundamental frequency \( f_0 \); 2) the average absolute deviation from the mean musical pitch \( \bar{\Delta} \) as an estimator of vibrato extent; and 3) the dominant modulation frequency \( f_{\text{mod}} \) (see further below).

Vibrato extent (i.e. the modulation amplitude) is typically expressed in semitones or cents, where one octave equals 12 semitones or 1,200 cents (11). For the purpose of this analysis, the frequency values of the fundamental frequency contours were converted to cents relative to middle C as

\[
c[i] = 1200 \log \left( \frac{f[i]}{f[\text{C4}]} \right) \log(2)
\]

where \( f[i] \) is the fundamental frequency estimate at time index \( i \), and \( f[\text{C4}] \) is the fundamental frequency of middle C (C4), calculated as \( 440 \times 2^{(-9/12)} \approx 261.63 \) Hz. The average musical pitch \( \bar{\tau} \) of the \( f_0 \) contour, expressed in cents from
middle C, was then computed as
\[ \zeta = \frac{1}{n} \sum_{i=0}^{n-1} \frac{c_i}{|c_i|} \]  (2)

Finally, the average absolute deviation from the mean musical pitch \( \Delta c \), representing vibrato extent, was computed as
\[ \Delta c = \frac{1}{n} \sum_{i=0}^{n-1} |c_i - \bar{c}| \]  (3)

As a reference, for completely regular vibrato with perfectly sinusoidal frequency modulation, the vibrato extent would equal twice the (maximum) modulation amplitude divided by \( \pi \), i.e. the amplitude divided by about 1.57. Consequently, a perfectly sinusoidal vibrato with a modulation amplitude of \( \pm 1 \) semitone would correspond to about \( \Delta c \approx 63.7 \) cents. In order to rule out analysis artifacts, four segments with \( \Delta c < 127.3 \) (representing a perfectly sinusoidal vibrato) and an extent of more than two semitones were excluded from the analyzed data set.

A completely regular vibrato naturally contains only one modulation frequency (see e.g. Figure 1A for an example of Luciano Pavarotti’s vocal vibrato with an almost sinusoidal modulation frequency of about 5.7 Hz—example 1.1 from Miller (12)). In contrast, preliminary inspection of Freddie Mercury’s vocal vibrato suggested more irregular frequency modulation patterns, caused by the superposition of more than one modulation frequency component (see Figure 1B for an example, taken from ‘Bohemian Rhapsody’, \( t = 71 \) s in The Acapella Collection release (\( t = 132 \) s in the A Night at the Opera album), ‘go’ within the phrase ‘Good-bye everybody, I’ve got to go’).

When considering the \( f_0 \) contour in Figure 1B it became apparent that time domain methods (like peak-picking) for determining the number of vibrato cycles per second are problematic: it is unclear whether the small ripples around \( t = 1 \) should be counted as individual cycles or not. Therefore, in order to obtain a robust numerical parameter for assessing Freddie Mercury’s vibrato, the dominant modulation frequency \( f_{\text{mod}} \) was computed with Fourier analysis. For each segment, the DC offset (i.e. the mean value) of each \( f_0 \) contour was removed. The normalized \( f_0 \) contour was scaled by a Hann window, and then zero-padded to eight times its original length. The resulting time series vectors were subjected to a forward FFT, and the amplitude spectrum was calculated. The frequency of the spectral maximum in the amplitude spectrum was considered as the dominant modulation frequency, \( f_{\text{mod}} \). In order to rule out artifacts, four segments with \( f_{\text{mod}} < 3 \) Hz were excluded from the analyzed data set.

As a by-product of the FFT analysis of the \( f_0 \) contours, the amplitude difference of the two strongest frequency components of the \( f_0 \) modulation contour was calculated as \( \Delta A_{f_{\text{mod}}} = (1 - \frac{A_1}{A_2})/A_1 \), where \( A_1 \) and \( A_2 \) are the two strongest components found in the normalized amplitude spectrum of the \( f_0 \) contour (dark triangles and light diamonds, respectively, in Figure 1A and B). The \( \Delta A_{f_{\text{mod}}} \) parameter is an estimator of the vibrato regularity. It can theoretically assume values in the range of 0 to 1, where 0 would represent purely sinusoidal frequency modulation with one modulation frequency component. When considering the preliminary analysis done in Figure 1, the respective \( \Delta A_{f_{\text{mod}}} \) values found in Luciano Pavarotti’s and Freddie Mercury’s vibrato are about 0.89 and 0.11, suggesting a quite regular vibrato for Luciano Pavarotti and an irregular vibrato with two almost equally strong modulation frequency components for Freddie Mercury.

**Imitation and analysis of rough voice production**

Preliminary perceptual analysis suggested that a portion of the available data material was sung with a pronounced degree of roughness. The phenomenon was most clearly audible in sustained notes found at \( t = 31 \) s, 39 s, 120 s (‘on’ within the phrase ‘turn it on’) in the a-cappella track ‘Let’s
Turn It On’ from Freddie Mercury: The Solo Collection (2000, EMI/Parlophone). The respective sound examples were extracted and played back to a professional rock singer (co-author D.Z.-B.) who served as the subject in previous studies (13,14). The singer was asked to imitate these sounds, and phonation was recorded with a 70°C14 rigid endoscope (Karl Storz, Tuttlingen, Germany) attached to a Hispec 1 b/w high-speed camera (Fastec Imaging, San Diego, CA, USA) which was operated at 4,132 frames per second, with a 300 W xenon light source (Richard Wolf GMBH, Knittlingen, Germany). The video recordings were stored on a computer as uncompressed AVI files. Digital kymograms (DKGs) (15,16) were extracted with a Jython plugin (http://homepage.univie.ac.at/christian.herbst/index.php?page=fiji) written by author C.T.H. for the Fiji image analysis framework (17).

The acoustic signal was recorded with a Sennheiser MKE-2 microphone (Sennheiser, Wennebostel, Germany) mounted on a headset at 15 cm distance from the singers’ mouth, using a custom-built microphone amplifier. The sound was digitized with a LSI LSI PC/C32 sound card (Loughborough Sound Images plc, Loughborough, Leicestershire, England) at a sampling frequency of 16,000 Hz. The digitized sound was stored in the smp file format using the software Soundswell and Phog (Electronix Hitech, Täby, Sweden), and was later converted to and stored as uncompressed wav.

Results

Speaking fundamental frequency

Fundamental frequency data extracted from the analyzed interviews was not normally distributed. Overall, an average speaking fundamental frequency of 130.1 Hz was computed, with a median at 117.3 Hz and the 5th and 95th percentiles at 89.1 Hz and 231.4 Hz, respectively. Boxplots of the fundamental frequency extracted from the six interviews are shown in Figure 2. The mean and median values are provided in Table 1. In order to test whether the interviewer had an effect on the speaking fundamental frequency, all data points were pooled by interviewer. Interviews conducted by Rudi Dolezal had a mean speaking fundamental frequency of 119.3 Hz, with the median at 111.2 Hz and the 5th and 95th percentiles at 87.3 and 174.5 Hz, respectively. The interviews

Table 1. Average and median speaking fundamental frequency values (including 5th and 95th percentile) for the six analyzed interviews.

<table>
<thead>
<tr>
<th>Year</th>
<th>Interviewer</th>
<th>Number of data points</th>
<th>Average f0 (Hz)</th>
<th>Median f0 (Hz)</th>
<th>5th percentile f0 (Hz)</th>
<th>95th percentile f0 (Hz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1984</td>
<td>Dolezal</td>
<td>26,124</td>
<td>119.2</td>
<td>111.1</td>
<td>87.1</td>
<td>175.9</td>
</tr>
<tr>
<td>1984</td>
<td>Wigg</td>
<td>39,487</td>
<td>142.2</td>
<td>125.6</td>
<td>96.5</td>
<td>253.4</td>
</tr>
<tr>
<td>1985</td>
<td>Wigg</td>
<td>11,795</td>
<td>130.4</td>
<td>121.5</td>
<td>93.7</td>
<td>187.3</td>
</tr>
<tr>
<td>1986</td>
<td>Wigg</td>
<td>22,772</td>
<td>120.9</td>
<td>107.2</td>
<td>86.2</td>
<td>218.7</td>
</tr>
<tr>
<td>1987</td>
<td>Dolezal</td>
<td>13,816</td>
<td>119.6</td>
<td>111.4</td>
<td>87.6</td>
<td>173.2</td>
</tr>
<tr>
<td>1987</td>
<td>Wigg</td>
<td>20,424</td>
<td>137.7</td>
<td>124.8</td>
<td>88.3</td>
<td>240.5</td>
</tr>
</tbody>
</table>

Figure 2. A: Interquartile range distributions of speaking fundamental frequency estimates for the six analyzed interviews. The whiskers indicate the 5th and the 95th percentile, respectively, and the superimposed stars indicate the group means. B: Histogram of all speaking fundamental frequency data, grouped by interviewer (Dolezal: dark gray; Wigg: light gray; Cumulative: white).
conducted by Andrew Wigg had generally higher speaking fundamental frequency values, with an average of 134.6 Hz, the median being at 120.2 Hz and the 5th and 95th percentiles at 90.0 Hz and 245.7 Hz, respectively. A Kruskal–Wallis rank sum test performed on the non-normally distributed data with the software R (18) revealed a highly significant difference between the two groups (Dolezal versus Wigg, $P < 0.001$).

**Singing voice range**

The lowest note found in the analyzed data corpus was a short F#2 (about 92.2 Hz) in 'Don’t Try Suicide' (*The Game*), found at $t = 52.5$ s (‘so’ within the phrase ‘So you think it’s the easy way out?’). A series of sustained notes at musical pitch G2 (about 98 Hz) were found in ‘Get Down Make Love’ at $t = 58.7$ s, 61.8 s, 116.7 s, 120 s, 146.7 s, and 149.9 s in the version from *The Acapella Collection* (‘down’ within the phrase ‘Ev’ry time I get hot you wanna cool down’) — note that in relation to the version on the *News of the World* album, these time offsets are shifted by about 12.5 seconds. For the highest notes a couple of candidates were found: 1) a sustained E5 (about 659.3 Hz) in ‘Hang on in There’ (*The Miracle*) at $t = 42$ s (background voice ‘yeah’, right after ‘majesty around the world’), which was masked by other sung sounds in the recording; 3) a very short but clearly perceivable G5 (about 784 Hz) in ‘Let’s Turn It On’ (*Freddie Mercury: The Solo Collection*) at $t = 25$ s (‘on’ within the phrase ‘let’s turn it on’), which was most probably produced in falsetto register.

**Descriptive analysis of functional singing voice parameters**

Overall, based on perceptual assessment, Freddie Mercury seemed to have ample control over vocal registration and ‘blending the registers’, i.e. mixing the chest and falsetto registers. More often than not, both these registers tended to have comparable timbral characteristics. Sounds produced in the falsetto register, where unambiguously identifiable (e.g. ‘Let’s Turn It On’ at $t = 23.2–25.2$ s), could contain harmonic content up to 10 kHz and sometimes even beyond. (Note, though, that any existing harmonic content could have been amplified by the sound engineer during the post-production stage of the recording.) A few examples where the choice of vocal registration was more evident due to abrupt transitions between the registers include: 1) ‘Nevermore’ (*Queen II*), $t = 15–28$ s (‘Can’t you see? Listen to the breeze, whisper to me please. Don’t send me to the path of nevermore’); 2) ‘Seaside Rendezvous’ (*A Night at the Opera*), $t = 34–45.2$ s (‘Can we do it again? Can we do it again sometime? (ooh I like that) Fantastic, c’est la vie Madames et Monsieurs’); 3) ‘Love of My Life’ (*A Night at the Opera*), $t = 131–134.6$ s (‘How I still love you’), where there was a perceivable and abrupt timbral shift between ‘love’ (sung in chest register) and ‘you’ (falsetto register); or 4) ‘Keep Passing the Open Windows’ (*The Works*) at $t = 18–23$ s, where ‘need’ of the phase ‘love is all you need’ again bore perceptual clues of falsetto voice production.

![Figure 3](https://example.com/figure3.png)

**Figure 3.** Examples of ‘breathy’ and ‘pressed’ sound production. A and B: narrow-band spectrograms (FFT bin width 10.77 Hz, dynamic range 55 dB); C and D: corresponding amplitude spectra, extracted at $t = 20.5$ s and $t = 2$ s, respectively. The light gray (orange color online) line indicates the noise floor at about −40 dB and −55 dB, respectively.
Freddie Mercury appeared to have had good control over the phonation type along the dimension ‘breathy’ to ‘pressed’. Two such examples illustrating the range of timbral possibilities are illustrated in Figure 3. The example for breathy phonation (Figure 3A and C), taken from ‘Teo Toriatte’ (A Day at the Races) contained clearly audible broadband noise above 1.5 kHz, which could hypothetically have been created by turbulent noise (19), presumably caused by incomplete vocal fold adduction. This broadband noise was strong enough clearly to highlight several vocal tract resonances, identified as the second, third, and fourth formant in the spectrogram (see arrow annotation in Figure 3A and C). An example for the ‘pressed’ phonation type is illustrated in Figure 3B and D. Apart from the perceptually strained quality of the sound and the apparent lack of vocal vibrato, the example was characterized by a clearly audible release of subglottal pressure at the end of the phrase (t = 5.2 s). When the normalized spectra of the two antagonistic examples are compared, clear differences of spectral composition become apparent. The ‘breathy’ example, although polluted by keyboard accompaniment, contained a far smaller number of noteworthy harmonics, and the normalized spectrum had a higher noise floor (about –40 dB overall) as compared to the ‘pressed’ example (about –55 dB)—see light gray (orange color online) lines in Figure 3C and D.

Vibrato

Boxplots of the calculated vibrato parameters are shown in Figure 4. The average fundamental frequency of the analyzed vibrato notes was 350.2 Hz, with a median of 347.0 Hz and 5th and 95th percentiles at 199.6 Hz and 494.9 Hz, respectively. The mean dominant modulation frequency $f_{\text{mod}}$ was 7.0 Hz (median 7.2 Hz, 5th and 95th percentiles at 5.2 Hz and 8.2 Hz). The average vibrato extent $\Delta c$ was calculated to be 35.0 cents (corresponding to a perfectly sinusoidal vibrato with an amplitude of about 55 cents or 0.55 semitones—see Methods), with a median of 32.1 cents and the 5th and 95th percentiles at 19.6 cents and 64.4 cents, respectively. The vibrato regularity estimator $\Delta A_{f_{\text{mod}}}$ had a mean value of 0.57, with a median at 0.60 and 5th and 95th percentiles at 0.09 and 0.97, respectively. The variation of dominant modulation frequency $f_{\text{mod}}$ and modulation extent $\sigma_c$ as a function of fundamental frequency were analyzed with linear regression analysis. While there was only a slight tendency for decreased dominant modulation frequency as $f_0$ increases ($f_{\text{mod}} = -0.0033 f_0 + 8.19, R^2 = 0.09$), there existed a slightly stronger (but still weak) inverse correlation between the $f_0$ and modulation extent ($\Delta c = -0.0674 f_0 + 58.98, R^2 = 0.21$). In other words, the vibrato extent tended to be smaller at higher pitches.

Subharmonic voice production

An example of Freddie Mercury’s phonation with a perceived degree of roughness is illustrated in Figure 5A. In the spectrogram, the harmonics appear as integer multiples of the fundamental frequency, and so-called subharmonics emerge at integer multiples of a third of the fundamental frequency, constituting a period-3 subharmonic pattern (period tripling). The same phenomenon is seen in the vocal imitation
produced by professional rock singer D.Z.-B. (Figure 5B). Analysis of the laryngeal endoscopic HSV data from D.Z.-B.’s phonation revealed frequency-locked vibration of the ventricular folds at a third of the frequency of vibration of the vocal folds (see Supplementary Movie M1, available online). This was also seen in the corresponding DKGs (Figure 6), extracted at two positions along the anterior–posterior vocal fold dimension. Anteriorly, the ventricular fold vibration exhibited clear lateral surface waves (similar to mucosal waves within the true vocal folds), possibly induced by high levels of glottal air flow and subglottal pressure. Overall, the epilaryngeal tube was constricted, thus limiting the visibility of the true vocal folds. The arytenoid cartilages (and thus also the vocal processes) were firmly adducted.

Discussion

The main difficulty of this retrospective study was posed by the fact that no controlled data acquisition was possible. The results of the acoustic and perceptual analyses conducted here need thus be treated with care: not all sound samples were a-cappella recordings, variable recording tape wheel speeds could have slightly changed the measured fundamental frequencies, and any spectral component could have been changed by the intervention of sound engineers at the post-production stage (therefore no formant data were analyzed). Except for the vocal imitation by rock singer D.Z.-B. no physiological data were available to back up the physiology-related conclusions drawn from the analyzed acoustic data. In spite of these complications, this study was successful in delivering a few insights into Freddie Mercury’s voice production and singing style. These are summarized and discussed below.

Speaking voice fundamental frequency

Analysis of interviews revealed a mean SFF of 130.1 Hz (median at 117.3 Hz). Morris et al. found an average SFF of 127.4 Hz (±10.1 Hz) for trained middle-aged baritones and basses (20); Brown et al. documented an average SFF of 119 Hz (no standard deviation given) for the same group (21); and Chernobel’skii reported a SFF of 123.2 Hz (±10.2 Hz) for baritones (22). Based on these data it can be hypothesized that Freddie Mercury had a baritone’s speaking voice, since the respective mean SFF values for trained tenors were about 20 Hz higher (20–22). This assumption is further motivated by the notion that the SFF is located at a roughly predictable offset within a speaker’s total f0 range (23). Normative values given for that offset in various studies are in the range of 3–8 semitones above the lowest achievable fundamental frequency (20,24,25), and this variability might be caused by effects of vocal training (23). Anecdotaly, the suggestion that Freddie Mercury had a baritone’s voice range is corroborated by a quote from his duet partner, classical soprano Montserrat Caballé, who maintained that ‘He had a baritone voice. I told him one day, “Let’s do a small duet of baritone and soprano”, and he said, ”No, no, my fans only know me as a rock singer and they will not recognise my voice if I sing in baritone”’ (2).

There was a noteworthy and statistically significant difference between SFF data from interviews by Andrew Wigg versus Rudi Dolezal. Interviews conducted by Andrew Wigg contained several spoken phrases where the SFF exceeded 200 Hz, with maximum values up to 320 Hz (which occurred during connected speech as a means of putting emphasis on certain words) as indicated by manual inspection of the raw data material. Other confounding factors notwithstanding, the SFF difference might be attributed to Andrew Wigg being a native speaker of British English (as was Freddie Mercury), thus creating a more familiar and personal atmosphere during the interviews that allowed for a somewhat higher SFF. In contrast, the Austrian Rudi Dolezal was a native speaker of German.

Singing voice range

Measurement of the singing voice range constituted the most challenging task in this study, since it was solely based on perceptual assessment of commercially available recordings. The determined extreme pitches of F#2 (about 92.5 Hz) and G5 (about 784 Hz) constitute a singing voice range of 37
reported to be possible for a ‘male soprano’ (27) and for while such high fundamental frequencies of notes were available, no rigorous proof can be given. The assumption made in this study versus laryngeal configurations (see Methods and Figure 1). In contrast, the analysis of Freddie Mercury’s notes sung with vibrato resulted in an average AFmod value of 0.57, and half the analyzed samples had values of 0.6 or smaller, indicating that largely more than one modulation frequency component was found in the spectral analysis of the F0 contours. The principal vibrato agent is assumed to be a muscle or a group of muscles that regulates fundamental frequency. Typically, the cricothyroid is hypothesized to form an agonist-antagonist pair with the thyroarytenoid and/or the lateral cricothyroid muscles (10, 38). In the ‘standard’ case of regular vibrato (see e.g. the Pavarotti example in Figure 1), such an agonist–antagonist system would exhibit simple sinusoidal oscillation. Such a simple model could, however, not explain the complex f0 modulation seen in some of Freddie Mercury’s sustained notes. It is thus hypothesized that either a single agonist–antagonist pair exhibited complex irregular oscillation, or more than one agonist–antagonist pairs for fundamental frequency modulation were present. Naturally, as no physiological data were available, neither position can be maintained. The proposed notion might, however, be relevant for further physiological investigation involving contemporary singers. Perceptually, Freddie Mercury’s irregular (and typically faster) vibrato is clearly audible in the sustained notes of famous songs such as ‘Bohemian Rhapsody’ (A Night at the Opera) or ‘We Are the Champions’ (News of the World), and it appears to be one of the hallmarks of his vocal style. On a more general level, these findings suggest that the regularity of vibrato (indicated here by the AFmod parameter) could be an important feature for any singing voice, which may deserve to receive further scientific attention. The analysis of rock singer D.Z.-B.’s imitation of Freddie Mercury’s ‘rough’ singing style revealed subharmonic vibration (39, 40) created by 3:1 frequency locking of the vocal folds and the ventricular folds: during each vibration of the ventricular folds the vocal folds completed three oscillations. Similar
vibratory patterns of the vocal and the ventricular folds as synchronized oscillators were observed in previous studies involving the same (14) and other singers (41,42). The observed lateral surface waves along the ventricular folds are comparable to mucosal waves in true vocal folds (43,44).

This type of voice production, involving vibrating ventricular folds at various degrees, has been termed ‘growl’ (45), ‘dist’ (14), ‘throat-singing’ (46), or ‘vocal-ventricular mode’ (VVM) (42). The umbrella term ‘distortion’ (47) or ‘intentional distortion’ (ID) (48) has been suggested for singing voice effects where, in addition to the ventricular folds, also other supraglottal structures are engaged in vibration.

Spectrograms of sound production by both D.Z.-B. and Freddie Mercury show comparable subharmonic patterns (period tripling). It is therefore likely that frequency-locked ventricular fold vibration also occurred in the original sample of Freddie Mercury’s ‘rough’ singing which was imitated in this study. An alternative, but less likely explanation would suggest subharmonic oscillation of the vocal folds alone (49), or a frequency-locked vibration of the vocal folds and the aryepiglottic folds (45).

**Voice production and stage persona**

Subharmonics, as described in the previous paragraph, are a possible route of a system ‘on its way to chaos’ (50), with potential communicative relevance (51). Their occurrence aids in creating the impression of a sound production system driven to its limits, even while used with great finesse. These traits, in combination with the fast and irregular vibrato, might have helped create Freddie Mercury’s eccentric and flamboyant stage persona. This stage persona and its apparent constant concourse with Freddie Mercury’s more private side is best characterized by some quotes from his interviews: ‘I’m just a musical prostitute’ (Dolezal, 1984); ‘I am the nicest person you could meet, my dear. It’s just that I am handicapped, and therefore I have to sort of virtuoso’ (Dolezal, 1984); ‘I’m just a musical prostitute’ (Dolezal, 1984); ‘I am the nicest person you could meet, my dear. It’s just that I am handicapped, and therefore I have to sort of virtuostically fight my so-called stage persona. Most of the times it sort of works against me’ (Wigg, 1984); and, concerning the audience’s expectations and losing control: ‘Because of my persona on stage they think that I carry on this way offstage. If I did, I would have been dead a long time ago’ (Wigg, 1984).

**Acknowledgements**

This study is a work of ‘fan-science’. Author contributions: C.T.H. conceived the study, collected the data material, conducted the perceptual and quantitative analysis, and wrote the manuscript. S.H., D.Z.-B. and P.A.L. collected the vocal imitation data and contributed to the final version of the manuscript. We are very grateful to Dr Hans Larsson (Karolinska University Hospital) for his help in acquiring the respective endoscopic data. Parts of this manuscript were previously published in: Christian T. Herbst (2012). Freddie Mercury—Akustische Stimm-Analyse. L.O.G.O.S. Interdisziplinär, 20 (3), 174–83.

**Declaration of interest**

The authors report no conflicts of interest.

**Funding information**

This study was partially financed by the institutional fund of the Palacký University Olomouc, Czech Republic (C.T.H.).

**References**
