JASA ARTICLE

Dynamic changes of vocal tract dimensions with sound pressure level during *messa di voce*[a)]

Fabian Burk,[1,2,b)] Louisa Traser,[2] Michael Burdumy,[3] Bernhard Richter,[2] and Matthias Echternach[4]

[1]*Department of Otorhinolaryngology and Plastic Surgery, SRH Wald-Klinikum Gera, Gera, Germany*

[2]*Institute of Musicians' Medicine, University Medical Center Freiburg, Freiburg im Breisgau, Germany*

[3]*Department of Radiology, Medical Physics, University Medical Center Freiburg, Freiburg im Breisgau, Germany*

[4]*Division of Phoniatrics and Pediatric Audiology, Department of Otorhinolaryngology, Munich University Hospital (LMU), Munich, Germany*

ABSTRACT:

The *messa di voce* (MdV), which consists of a continuous crescendo and subsequent decrescendo on one pitch is one of the more difficult exercises of the technical repertoire of Western classical singing. With rising lung pressure, regulatory adjustments both on the level of the glottis and the vocal tract are required to keep the pitch stable. The dynamic changes of vocal tract dimensions with the bidirectional variation of sound pressure level (SPL) during MdV were analyzed by two-dimensional real-time magnetic resonance imaging (25 frames/s) and synchronous audio recordings in 12 professional singer subjects. Close associations in the respective articulatory kinetics were found between SPL and lip opening, jaw opening, pharynx width, uvula elevation, and vertical larynx position. However, changes in vocal tract dimensions during plateaus of SPL suggest that perceived loudness could have been varied beyond the dimension of SPL. Further multimodal investigation, including the analysis of sound spectra, is needed for a better understanding of the role of vocal tract resonances in the control of vocal loudness in human phonation.
 https://doi.org/10.1121/10.0022582

(Received 12 July 2023; revised 11 October 2023; accepted 14 November 2023; published online 1 December 2023)

[Editor: Zhaoyan Zhang] Pages: 3595–3603

I. INTRODUCTION

The control of vocal loudness can be considered a key element for successful communication in different emotional and environmental circumstances.[1] The extent to which loudness can be varied—in addition to the variation of pitch—is considered an important factor for determining the health of a voice, e.g., in the voice range profile.[2] It certainly is important in singing, where the variation of sound volume—referred to as dynamics—plays a major role.

A special kind of vocal loudness variation is called for in the so-called *messa di voce* (MdV). It consists of a continuous crescendo and decrescendo of a tone without changing its pitch and is considered a central technique in the technical repertoire of Western classical singing.[3] Since the increase in loudness by increasing the subglottic pressure is physiologically accompanied by an increase in the fundamental frequency (f_o),[4,5] a complex interaction between the adaptation of the lung pressure and adjustments on the laryngeal and vocal tract level is needed[3,6] to meet the requirements of MdV.

The trias of principle ways to influence vocal loudness is well established.[4,7–9] The first component is the variation of lung pressure as the primary energy source of the oscillating system of the vocal folds. Second, glottal adduction and vertical vocal fold thickness are—amongst others—major influence factors on loudness on the glottal level. The third factor is the variation of the shape of the vocal tract, defined as the air containing structures from above the glottis to the lips. By passing the vocal tract, the source spectrum is altered by the resonant properties of the vocal tract into the final product, namely, the sound, which is radiated from the lips. These resonant properties can be changed by articulation, that is, by movement of the mobile components, which define the shape of the vocal tract like tongue, velum, jaws, and lips.[10–12] Titze[8] showed in simulations that certain variations of the vocal tract alone could increase vocal loudness without changing lung pressure or glottal adduction.

Multiple forms of adjustments of vocal tract dimensions in changing sound pressure level (SPL) and/or loudness conditions have been described in the literature. Vocal tract resonances can be tuned to strong harmonics in the source spectrum to gain high values of SPL, which is mostly determined by the harmonic nearest to the first vocal tract resonance (f_{R1}).[4,7] In a special scenario, this can be exploited by female voices in the f_o range under 750 Hz—or even higher[13]—by tuning f_{R1} to f_o.[14–16] Another option, which is frequently used by male voices in Western operatic singing, is so-called resonance clustering, where the clustering of $R_{3–5}$ results in spectral peak around 3 kHz. This peak has been termed the singer's formant cluster and falls into an

[a)]Portions of this work were presented in "[Messa di laringe - the dynamic relationship between larynx position and loudness in messa di voce]" at the 39th scientific conference of the German Society of Phoniatrics and Pediatric Audiology in Cologne, Germany, September 2023.

[b)]Email: fabian.burk@gmx.de

area where the human auditory system is very sensitive, so that a gain in perceived loudness is achieved.[7,17] Changes of vocal tract dimensions also affect physiological parameters related to SPL and loudness on a glottal level by influencing the flow pulse[18] or assisting the vocal fold vibration via an increase in vocal tract inertance.[19] According to Sundberg, the lowering of the larynx physiologically leads to some degree of vocal fold abduction, which should facilitate flow phonation with positive effects for vocal loudness.[4]

MRI has proven to be an appropriate method within voice research to investigate the dimensions of the vocal tract under different phonatory conditions.[13,20–23] Studies on vocal tract dimensions in different loudness conditions that consider Western classical singing are limited; however, some data are available. Ikävalko *et al.*[20] investigated loud singing in three female singer subjects in different singing styles and found loud singing to be achieved with very different vocal tract morphologies in the three studied singing styles kulning, edge, and operatic singing. The latter was characterized by a comparatively low vertical laryngeal position, large tongue-palate distance, and large pharyngeal diameters. Mainka *et al.*[21] studied five female and five male Western-classically trained singers with static magnetic resonance imaging (MRI) and found for high SPL values a comparatively low vertical larynx position, small measures concerning the epilarynx tube, and a reduction of the volume ratio epilarynx tube/hypopharynx, but with some differences between the sexes. In 2016, our research group published a dynamic real-time MRI study, where 12 professional classical singer subjects were asked to sing ascending scales in an MRI scanner in three different loudness conditions (*mezzo forte*, *pianissimo*, and *fortissimo*).[24] Different articulatory parameters were measured and studied concerning their correlation with SPL and musical dynamic level. This yielded a variety of significant correlations, namely, for the parameters lip opening (LO), jaw opening (JO), uvula elevation (UE), pharynx width (PW), and larynx position (LP). A limitation concerning the interpretation of this study was the fact that there were also multiple correlations between articulatory parameters and musical pitch and between SPL and pitch. Thus, pitch acted as a confounder variable, and it was not possible to cleanly separate the effects of SPL and pitch on articulation. Furthermore, some of the used scales incorporated a region of pitch where registration events are expected to occur,[25–27] so that effects of registration on articulation could not be ruled out.

Vocal tract adjustments attributable to continuous variation of loudness have not—to our knowledge—been studied in detail. Thus, this study aims to analyze dynamic changes in vocal tract dimensions of Western classical singers associated with continuous changes of musical dynamic level. To represent the latter, SPL had to be chosen as a surrogate parameter. To be able to study the effect of changes in SPL on articulation, the MdV task was performed in one vocal register only. The parameters that had shown significant correlation with the loudness condition in our previous study[24] were selected for further investigation in the present study.

II. MATERIAL, STUDY POPULATION, AND METHODS

The study was approved by the local ethics committee. Twelve professionally trained Western operatic singer subjects (three sopranos, three mezzo-sopranos, three tenors, and three baritones) were included after having given their written consent. All subjects are well established in their respective professional fields on a national or international level, as can be seen in Table I. When asked about their current voice condition, none complained of any disorders. A clinical examination by videostroboscopy and/or high-speed digital imaging was performed before the experiment and showed no pathologies. The general conditions of the experiment and the study group are largely the same as described in a previous study.[24]

All subjects were asked to perform a MdV on the vowel [aː] and on the pitch D4 (corresponding to an f_o of 294 Hz) while lying in supine position in the MRI scanner [3.0 T TIM TRIO (Siemens, Munich, Germany)]. The task was to be performed in chest voice (modal register) without register changes to rule out registration as a confounder variable with known influence on articulation.[25,29] Only tasks that were considered as properly executed by both the investigators (amongst whom were two ENT specialists and singers) and the subject were further investigated. Only one task per subject was analyzed.

The MRI images were recorded by 2D radial flash sequence[30] with approximately 25 frames/s in the midsagittal plane (resolution 128×128 pixels, field of view

TABLE I. Subjects, voice classification, and taxonomy according to Bunch and Chapman (Ref. 28).

Subject	Classification	Taxonomy
1	Soprano	2.1/2.4—International opera principal/ international concert-oratorio
2	Soprano	2.1/2.4—International opera principal/ international concert-oratorio
3	Soprano	3.15b1/3.17—Professional chorister adult/ensemble
4	Mezzo	2.1—International opera principal
5	Mezzo	2.1/2.4—International opera principal/ international concert-oratorio
6	Mezzo	2.4—International concert-oratorio
7	Tenor	2.4/2.1—International concert-oratorio/ international opera principal
8	Tenor	2.1/2.4—International opera principal/ international concert-oratorio
9	Tenor	2.4/3.1a—International concert-oratorio/ national major principal opera
10	Baritone	2.1/2.4—International opera principal/ international concert-oratorio
11	Baritone	2.4/3.1a—International concert-oratorio/ national major principal opera
12	Baritone	2.4/3.1a—International concert-oratorio/ national major principal opera

$210 \times 210 \times 10\,\mathrm{mm}^3$). Synchronous audio recordings were taken with an optical microphone system (CONFON HP-SI 01, MR confon GmbH, Magdeburg, Germany), and audio feedback was provided to the subjects via headphones. The audio signal underwent a two-step filtering process: Directly upon recording, the scanner noise was automatically reduced by dedicated software built into the Confon system. Prior to further analysis, the scanner noise was then selected as a reference signal in Adobe Audition before the start of phonation in the recording of each subject and filtered out throughout the recording using the software's own adaptive filter. The microphone and filter settings are described in more detail in the aforementioned study.[24] The SPL of the audio files was analyzed with Praat[31] via the function "intensity listing." Instantaneous linear SPL values (i.e., without frequency or time weighting) were computed with a time step of 8 ms. Then each frame of the MRI recording was assigned the median SPL value of its corresponding time window. Due to the magnetic environment and the confined conditions, including the presence of a head coil, the mouth-to-microphone distance of the headset was not fixed, and the SPL values were not calibrated. However, these same cramped conditions also only allowed a variation of a few centimeters in the mouth-to-microphone distance between subjects.

After acquiring the MRI data, the resolution of each frame was set to 2048×2048 pixels (bicubic interpolation) for measuring purposes, and measurements of vocal tract dimensions were taken (see Fig. 1). Five articulatory parameters, LO, JO, UE, PW, and LP, which had shown

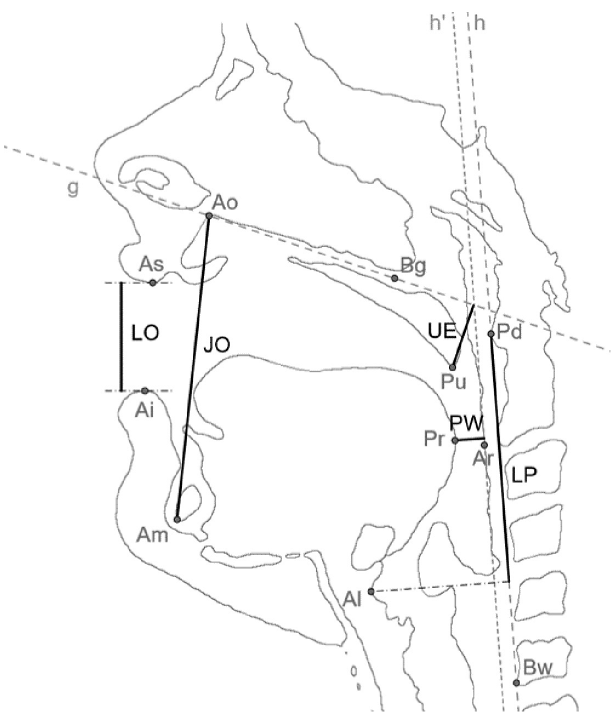

FIG. 1. Synopsis of the measured articulatory distances (solid black lines) in each frame. The names of the articulatory parameters are given in Table II. The anatomical landmarks and resulting auxiliary lines (gray) are explained and defined in detail in previous publications (Refs. 24 and 32).

significant correlations with SPL and/or the dynamic level in our previous study,[24] have been chosen for further investigation. Although vocal tract dimensions related to the epilarynx tube and the laryngopharyngeal junction have been shown to be associated with vocal loudness in the literature,[21] they could not be meaningfully evaluated in the real-time MRI images due to reduced spatial resolution in this anatomical area.

A synopsis of the measured distances is given in Fig. 1 and Table II.

For better comparability, all measured data were smoothed by centred moving average (window length 25 frames) and scaled from 0 (minimum value) to 1 (maximum value). To enable the comparison of the respective articulatory parameters as average over all subjects in one graph despite the different number of data points in the time series in different subjects, the data of each subject were interpolated to 100 equally spaced points of time (0.01–1) by linear interpolation (least square method). Figure 2 gives an overview of the described post-processing of the measured data.

For the creation of box plots, the webtool BoxPlotR[33] was used. The extent of the whiskers is computed by the Tukey method. To determine the statistical significance of differences, a two-sided, paired-samples t-test was performed. The significance level was set to $\alpha = 0.05$.

For the articulatory parameters, the neutral term vertex was chosen rather than maximum, because (1) in some cases the articulatory parameters showed contrary courses compared to SPL, so that a minimum in the articulator corresponded to a maximum in SPL [cf. the data for UE and SPL; see Fig. 4(a)], and (2) in some cases, the articulatory data only showed an association with SPL over a large part of the task but reached extrema at the beginning or end of it, so that a relative extreme of the articulator corresponded to the maximum value of SPL (SPL_{max}) (cf. the data for UE and SPL for S12 in supplementary material Fig. 2).

Due to the small number of subjects, higher grade regression analysis was not performed to investigate the

TABLE II. The measured articulatory parameters with their respective definitions.

Abbreviation	Articulatory parameter	Description
LO	Lip opening	Vertical distance between the lips.
JO	Jaw opening	Distance between the *spina nasalis* anterior and the most anterior-inferior point of the bone marrow in the *corpus mandibulae*.
UE	Uvula elevation	Distance between the tip of the uvula and the auxiliary line *g*, which represents the maxilla.
PW	Pharynx width	Smallest distance between the back of the tongue and the posterior wall of the oropharynx.
LP	Larynx position	Distance between the *tuberculum anterius atlantis* and the rectangular projection of the *antior commissura* on the line *h* representing the axial direction of the cervical spine.

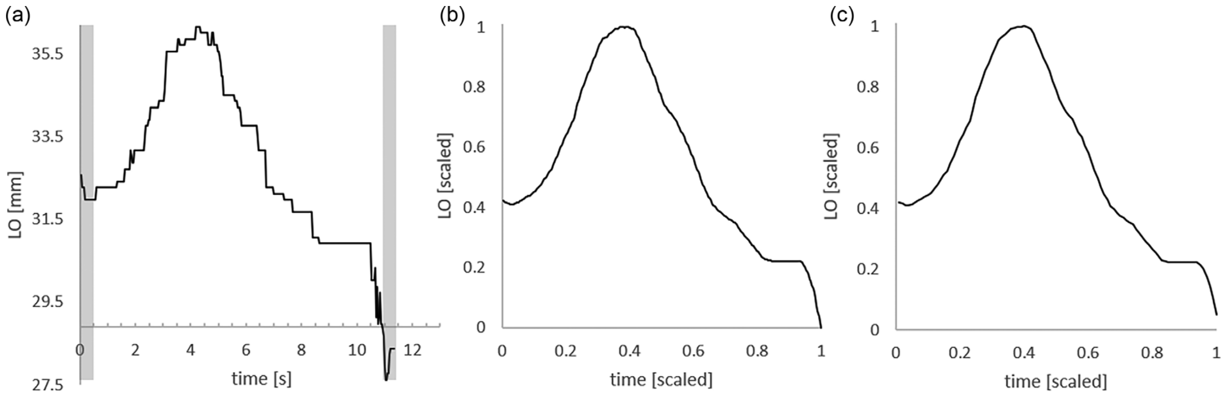

FIG. 2. (a) Original data of one subject for LO in mm vs time in seconds. (b) Smoothed data scaled from 0 to 1. Note that the first and last 12 data points are missing due to the smoothing process [gray areas in (a)]. (c) Interpolated data to 100 equally spaced time points from 0.01 to 1.

dynamic correlation between articulation and SPL. Concerning this relationship, the somewhat arbitrary term *association* was therefore used. The articulatory data were considered to be associated with SPL, when (1) they showed a clear vertex not further away from the time point of maximum SPL than 20% in scaled time, (2) they exhibited a traceably parallel or contrary course to SPL in this area, and (3) no oscillation of the graph of more than 20% value that was not reflected in the SPL graph occurred in this area.

III. RESULTS

The range of intensity of the MdV was 24.65 dB with a standard deviation (SD) of 6.1 dB over all subjects. There was no statistically significant difference between female and male singers, although a trend to a higher intensity range was visible for the male subjects [cf. Fig. 3(a)].

All subjects but one (S1) showed multiple associations of the measured articulatory parameters with SPL over time. S1 met the criteria for association with SPL for the parameter UE; the course of the other parameters was seemingly unrelated to SPL (cf. S1 in supplementary material Fig. 1).

On the level of the study population, all articulatory parameters showed visible associations with SPL in their kinetics over time. Furthermore, near SPL_{max}, the articulatory parameters tended to reach their respective vertices. For LO and JO, these vertices seem to be somewhat delayed compared to SPL_{max} [cf. S6; Fig. 3(b)].

The crescendo part was significantly shorter than the decrescendo part (means = 0.413 vs 0.587 in scaled time; $p = 0.011$).

Even when articulatory mean values of all subjects—including those who did not meet the criteria for association in single parameters—were compared to mean values of SPL, clear trends of unidirectional—i.e., positive (LO, JO, LP, PW) or contrary, i.e., negative (UE)—association with SPL were visible over time (cf. Fig. 4).

For the parameter LO, all subjects but S1 showed visible associations with SPL over time. There were significant differences between the means of LO at 50% SPL in the crescendo part (LO_{C50}) and LO at SPL_{max} (LO_{SPLmax}), as well as LO_{SPLmax} and LO at 50% SPL in the decrescendo part (LO_{D50}) of the task (both $p < 0.001$). Overall, LO showed a positive, i.e., unidirectional, association with SPL over time. For the decrescendo, the association was more pronounced than for the crescendo.

Concerning JO, the criteria for association with SPL were met in 9 of 12 subjects, excluding S1, S8, and S10.

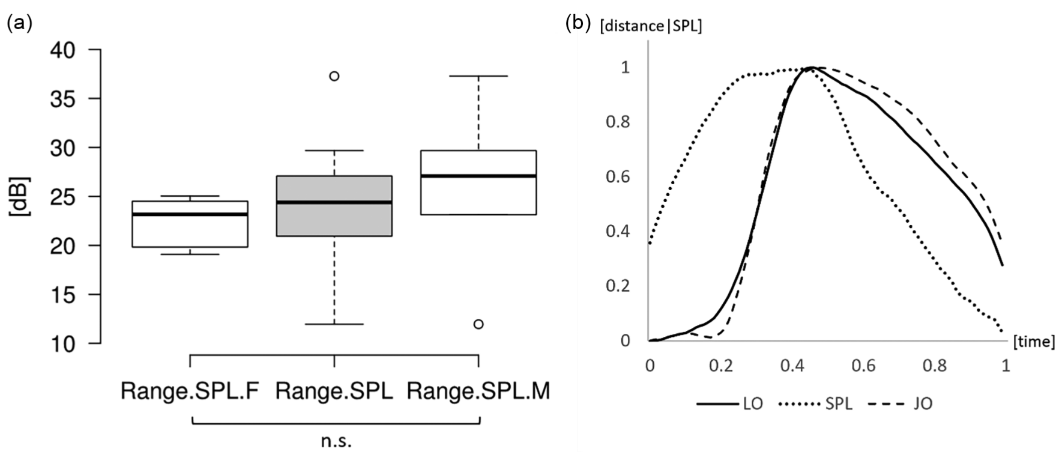

FIG. 3. (a) SPL ranges for female subjects (F), overall, and male subjects (M). n.s., not significant. (b) Course of LO, JO, and SPL in S6. Note the delayed vertex of the articulators after the SPL plateau is reached.

JASA
https://doi.org/10.1121/10.0022582

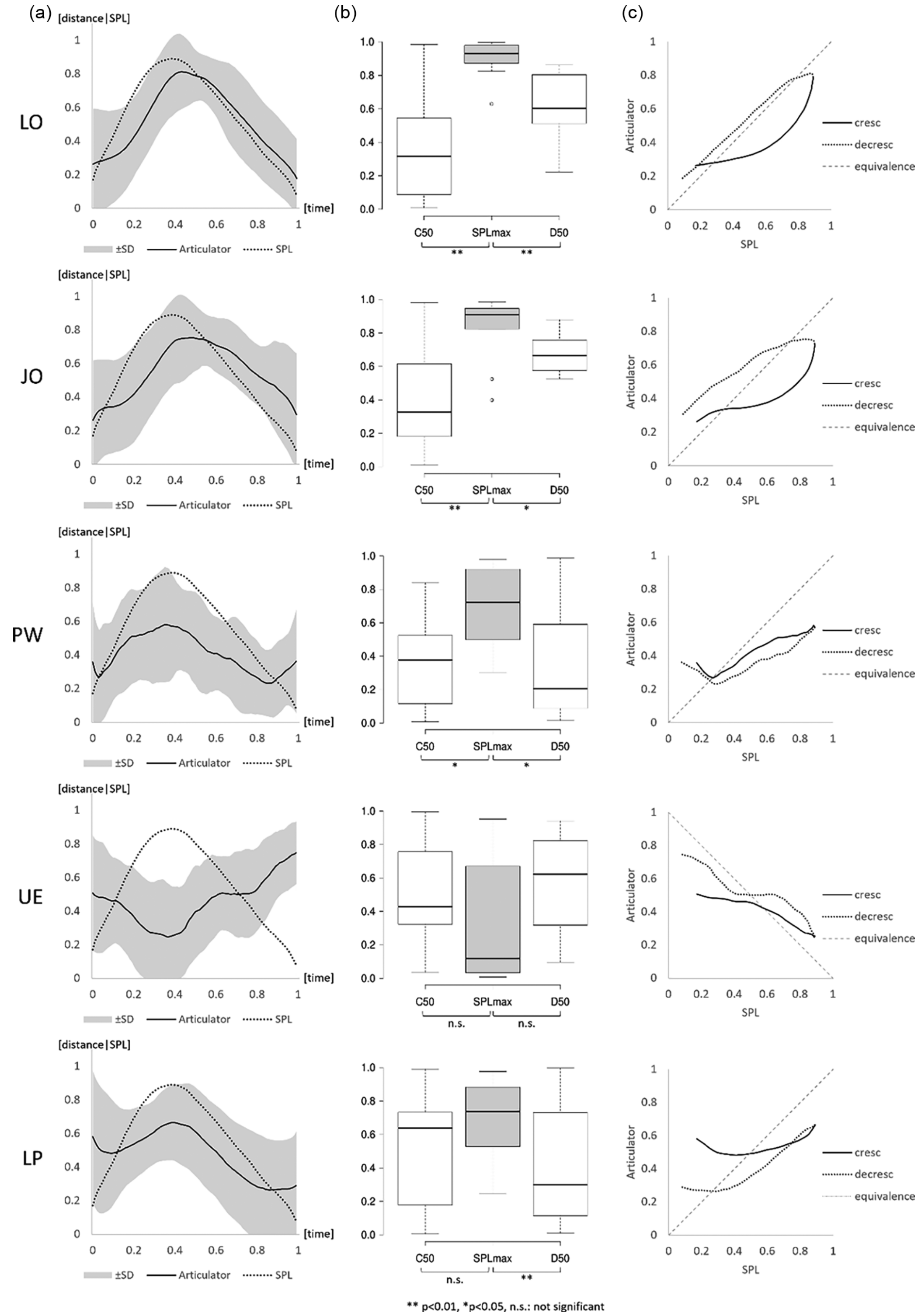

FIG. 4. (a) Course of the mean values of the respective articulators over all subjects, with the SD given in grays. The dotted line shows the course of the mean values of SPL over all subjects. Note that greater values for LP mean a lower larynx. (b) Variation of the articulatory values over all subjects at 50% SPL during crescendo (C50) and decrescendo (D50) and SPL_{max}, respectively. (c) The respective articulators plotted against SPL (mean values over all subjects).

JO_{SPLmax} was significantly greater than JO_{C50} and JO_{D50} ($p < 0.001$ and $p = 0.016$) over all subjects. In general, there was a unidirectional association with SPL over time. This parameter also followed SPL more closely in the decrescendo part than in the crescendo part, where SPL showed consistent growth from the beginning, while the articulatory parameter lagged somewhat behind at first and then reached its vertex almost simultaneously after a considerable increase in gradient.

The quotient LO/JO showed a visible association with SPL in itself (cf. Fig. 5), meaning that with rising and falling SPL, there was an additional kinetic in LO per JO.

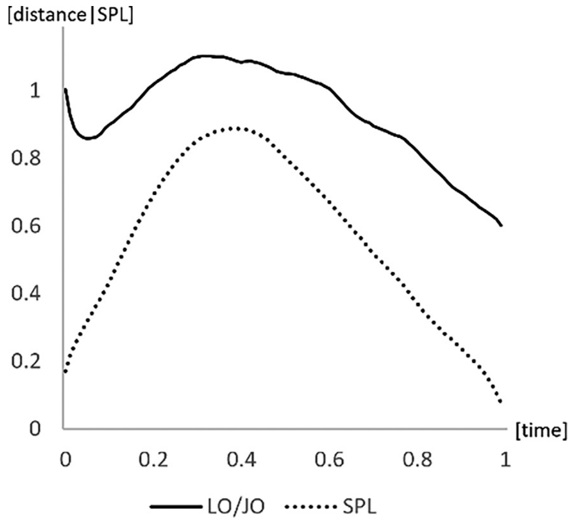

FIG. 5. Course of LO/JO compared to SPL over time (mean values over all subjects).

Only 6 of 12 subjects met the criteria for association of PW with SPL (S2, S3, S5, S7, S9, and S11). As in LO and JO, over all subjects, there was a significant difference of PW_{C50} and PW_{D50} to PW_{SPLmax}, respectively ($p = 0.016$ and $p = 0.013$) in the study population. The overall association of the articulatory parameter to SPL was unidirectional over time.

For the LP, seven subjects showed a clear unidirectional association to SPL (S2, S3, S5, S7, S8, S10, and S12). That means a lower larynx for louder conditions. Additionally, S9 presented a kind of inverse association, albeit with a vertex out of the 20% time corridor of tmax.SPL. S11 is also worth mentioning here, because there seems to be a unidirectional association between LP and SPL for the decrescendo part, but not for the crescendo part (cf. LP and SPL for S11 in supplementary material Fig. 2). The crescendo part exhibits almost no movement at all for this parameter, and a clear vertex is missing. Again, there is significant difference in the values for LP at SPL_{max} and D50 ($p = 0.008$) overall but not between C50 and SPL_{max} ($p = 0.214$).

The results for UE were somewhat heterogeneous. All subjects but S6 showed an association of UE and SPL. For S1, S2, S3, S5, S7, S9, S10, and S11, this association was of inverse nature, which means that the uvula was higher for greater values of SPL and lower for smaller ones. Conversely, S4, S8, and S12 exhibited a unidirectional association with SPL, which means a lower tip of the uvula for greater values of SPL. The overall trend of association seems to be inverse, as is shown in Fig. 4(a). Considering the two contrary types of association, the differences between UE_{C50}/UE_{D50} and UE_{SPLmax} did not reach the level of significance ($p = 0.232$ for the crescendo and $p = 0.144$ for the decrescendo). Furthermore, three subjects let the tip of the uvula sink to an extent that there was a velopharyngeal opening between oro- and nasopharynx visible in the midsagittal plane. While S3 exhibited this behavior only for a short moment at the very end of decrescendo, S11 performed the larger part of the task with an open velopharyngeal space closing it only between ∼80% and 30% decrescendo. S8 did not show any velopharyngeal closure over the whole task at all.

Furthermore, the behavior of the articulatory parameters seemed to be different in the crescendo part compared to the decrescendo part in all analyzed parameters, most markedly for LO, JO, and LP [cf. Fig. 4(c)].

IV. DISCUSSION

The present study analyses articulatory movement and SPL in real-time continuously for each point in time during MdV in professional Western classical singers. Close associations in the respective kinetics were found between SPL and LO, JO, PW, UE, and LP. Averaged over all subjects, LO, JO, PW, and LP show a unidirectional course with SPL, whereas UE shows inverse kinetics. However, there were considerable interindividual differences in articulatory movements.

Although anatomically closely linked, LO showed an additional kinetic per JO in association with SPL, as reflected by the fact that the LO/JO ratio itself was associated with SPL in a unidirectional manner (cf. Fig. 5). This is consistent with prior findings,[24] as is the increase in LO, JO, and PW for louder compared to softer phonation. Conversely, Mainka *et al.*[21] found no characteristic behavior for JO and an oropharynx measure comparable to PW with respect to SPL. The combination of the increase in LO, JO, and PW with rising SPL leads to a more megaphone-shaped vocal tract. An increase in JO has been shown by Lindblom and Sundberg[12] to raise the first resonance frequency of the vocal tract, and the additional lip opening per jaw opening for the louder condition should result in an acoustic shortening of the vocal tract, thus, uniformly raising its resonance frequencies. A wide LO in combination with a shortened vocal tract has been connected to a twangy sound quality, albeit in a speech study.[23] Pharynx width is mainly determined by tongue body position, and an increase in PW should lead to a markedly raised f_{R2}.[12] The net effect of the observed articulatory changes in LO, JO, and PW with increasing SPL could, therefore, be assumed to lead to a brighter sound quality, as respectively higher spectral components are boosted in the radiated spectrum.

Some subjects (cf. S6, S7, S10, and S11; supplementary material Figs. 1 and 2) reached a kind of SPL plateau in the middle of the MdV, whereas the articulatory parameters, especially LO and JO, continued to increase [cf. Fig. 3(b)]. This could indicate that the crescendo was continued beyond the objective category of the SPL on the level of loudness perception—by modification of the radiated spectrum. Indeed, Titze *et al.*[3] noted that some singer subjects who exhibited such a plateau of the SPL curve during MdV probably achieved an increase in perceived loudness through higher spectral content and vibrato extent. Similarly, Collyer *et al.*[34] could show an increase in higher spectral content toward SPL_{max} during MdV. The delay of the vertex

https://doi.org/10.1121/10.0022582

of LO and JO compared to SPL in the overall average of all subjects could be an expression of the presence of this strategy. Due to the scanner noise and the necessary filtering process of the audio recordings, the spectral properties could not be analyzed in a meaningful way.

The majority of the subjects performed the MdV with a closed velopharyngeal opening during the whole task, and 8 of 12 singers elevated the tip of the uvula for higher values of SPL. This behavior is in agreement with the results of our previous study.[24] Likewise, an inverse correlation for UE and SPL was found in some singers. A lowering of the tip of the uvula for higher SPL values occurred in a mezzo, a tenor, and a baritone in the present study. S8 (a tenor), did not close the velopharyngeal port during the whole task. This could be part of the personal singing technique. An open velopharyngeal port leads to nasalance. In literature, the occurrence of nasalance in classical singing has been described for the fine tuning of vocal timbre[35–37] and in the male passaggio region,[35,38] as well as being correlated to dynamic level.[39] It should be noted that nasalance was not measured in our study and that information on velum behavior is limited to the midsagittal plane. Thus, an acoustic coupling of the nasal tract to the VT could occur, even when there is no visible opening in the midsagittal plane. The fact that an open velopharyngeal port was visible at the end in S3 toward late decrescendo and S11 only closed it during parts of the decrescendo does not add any systematic insight, so that the use of nasalance during MdV remains unclear.

A trend to lower the larynx during crescendo and to raise it during decrescendo was observed over all subjects but not in every individual. A lower larynx for louder singing conditions has been previously observed in multiple studies concerning Western operatic style[21,24,40] and some non-classical styles.[41] The acoustic and vocal implications of a lowered larynx have been outlined as follows. (1) A lower larynx leads to a longer vocal tract, thus, uniformly lowering its resonance frequencies. (2) The mild abducent effect supposed to be associated with a low larynx position could also facilitate flow phonation with positive effects for loudness.[4,42] (3) A low larynx is considered to be important for the narrowing of the epilarynx tube and a low larynx to pharynx area ratio to produce the singer's formant cluster.[17,19,21,43] According to Titze,[19] the low larynx with narrow epilarynx tube raises the inertance of the vocal tract, thus, facilitating self-sustained oscillation of the vocal folds and improving glottal efficiency. The same author could show in simulations[8] that this vocal tract shape reduces spectral slope, boosts higher spectral content in the radiated spectrum, and therefore contributes greatly to increased loudness perception. However, as Mainka *et al.*[21] observed, a low larynx to pharynx ratio can also be achieved with high larynx positions, which could offer an explanation for the fact that there was no, or no clear, unidirectional association of LP and SPL in some subjects in our study (cf. especially the tenors S7 and S9 in supplementary material Fig. 2). In non-classical styles, a raised larynx has been linked to a more "metallic" sound.[22] Although the application of the singer's formant cluster to boost high spectral content was mostly observed in male voices and altos,[4,21] and not in sopranos, it should be taken into account that all subjects performed the MdV on the same pitch D4 and in chest voice. This suggests that the use of this strategy to increase perceived loudness would have been possible for the sopranos in our study too. Unfortunately, the immediate supralaryngeal dimensions could not be evaluated in our study due to a lack of spatial resolution of the dynamic MRI sequence.

The observed asymmetries are worth considering. In our study, the crescendo part was significantly shorter than the decrescendo part. In the aforementioned MdV study by Titze *et al.*,[3] the most experienced singers who produced the greatest SPL range exhibited asymmetries with longer crescendo and shortened decrescendo. This is in contrast to our findings. In fact, only two subjects (S4 and S10) performed the task with a longer crescendo than decrescendo: In the case of S4, about 40% of the task was used to reach 85% of SPL_{max} followed by 20% of slower SPL increase until SPL_{max} was reached. Interestingly, almost the entire articulatory increase in LO and JO happened in that 20% of the time [compare to the articulatory kinetics in S6; Fig. 4(c)]. With an SPL range of 25 dB, S4 was in the high-range region compared to the other females [cf. Fig. 3(a)]. S10 presented an almost symmetrical task, where the crescendo was only 2% longer than the decrescendo with an SPL range of 25 dB, ranking among the average values for male voices. All our subjects can be classified as very experienced professional singers.

Concerning the association of the articulatory parameters with SPL during crescendo and decrescendo, it can be stated that in the presented dataset, this association seemed closer during decrescendo than during crescendo, at least for the parameters LO, JO, and LP [cf. Fig. 4(c)]. To vary vocal loudness, it is assumed that three partly interdependent strategies[8,9,44,45] are applied: the variation of lung pressure, glottal adjustments, and the shape of the vocal tract. Although a MdV consists of a bidirectional variation of loudness, the variation of lung pressure is strongly dependent on the unidirectional lung volume. In addition, the provision of the lung pressure in the first section of the emptying of the lung volume up to the resting expiratory level is a rather passive process, partly even moderated by inhalatory muscles, whereas the pressure afterward has to be built up actively by muscle power (cf. the overview in chapter 3 of Sundberg's monograph[4] and the MRI-based analysis of respiratory kinematics during sustained phonation[46]). Thus, the precondition of the three-factor system changes over time, which could contribute to the almost hysteresis-like shapes of the articulator curves when seen as a function of SPL [cf. Fig. 4(c)]. It could be assumed that the extent to which each of the three factors contributes to the crescendo-decrescendo shape is dependent on the time point of interest. The respective lung volume even has an effect on the vocal tract shape independent of the SPL insofar as larger lung volumes tend to lead to a lower larynx.[47] This could explain the findings of higher

LP values at the beginning of crescendo compared to the end of decrescendo as well as the seemingly loose association of the course of LP to SPL during crescendo and the fact that the difference in LP between 50% SPL and SPL_{max} did not reach the level of significance during crescendo, whereas this was the case for decrescendo. Thus, it would seem that the effects of lung volume on LP mask those attributable to SPL change.

A. Limitations

This study focuses only on the Western classical tradition of singing. Thus, its results might not be representative outside of this genre, as vocal tract morphology has been shown to vary considerably between singing styles.[20,22] Although measures were taken to make the results comparable among subjects by standardized conditions and scaling of the values from 0 to 1, the generalizability of the results is likely to be limited due to the relatively small study population and high intersubject variability. As only one task per subject was analyzed, no information on possible intrasubject variability in the performance of the task is available. In addition, the conditions under which the MdV task was to be performed were quite artificial not only because of the MRI conditions—scanner noise, supine position etc.—which were discussed in detail in a previous study,[24] but also because of the instruction that they were to be performed entirely in the modal register. During a more natural phonation in classical singing, registration effects are likely to play a role in loudness control during MdV and, thus, could be expected to reflect on articulation.[25–27] Also, different vowel conditions lead to different vocal tract shapes, so it is possible that different adaptation strategies could be expected in different vowel conditions. As pointed out, the SPL measurement could not be calibrated, and the mouth-to-microphone distance varied on the order of a few centimeters. Thus, absolute SPL values have to be interpreted with caution. However, this affects the comparability of ranges only to a lesser extent and does not affect the course of SPL within a participant and the comparability of the scaled values among different subjects at all. Due to the scanner noise and the necessary filtering process, reliable spectral data of the audio signal were not available, and loudness in sones could not be determined. Yet a singer's crescendo and decrescendo will naturally aim to achieve an increase-decrease effect in perceived loudness in the ear of the listener, not in the physical metric of sound pressure. Thus, SPL is only a limited surrogate marker for the actually intended loudness, and it is very likely that the measures taken by the singer—be it adjustment of lung pressure, at the glottal level or considering articulation—are not fully reflected in SPL. The lagging kinetics of LO and JO during an SPL plateau in some subjects as discussed above could be an indication of this. Additional aspects like vibrato characteristics, which could not be assessed, are likely to contribute to the psychoacoustic increase-decrease effect inherent to MdV.[3,48]

V. CONCLUSIONS

Adjusting the shape of the vocal tract plays a role in classical singing in achieving increases and decreases in loudness during MdV. These adjustments are to some extent path-dependent, so that the return path during decrescendo does not seem to just be a mirrored image of the outgoing path during crescendo. The presented data point to the possibility that perceived loudness is increased beyond the dimension of SPL by vocal tract modifications that boost higher spectral content. In future studies, dynamic measurements in three dimensions that allow the assessment of all vocal tract sections—including the epilarynx tube and the nasal cavities—are desirable. Ideally, these could be matched to parallel information on lung pressure, glottal adjustments, and spectral data as well as psychoacoustic measures.

SUPPLEMENTARY MATERIAL

See the supplementary material for a graphical synopsis of the course of the articulatory parameters and SPL over time for each subject (SuppPub1.pdf).

ACKNOWLEDGMENTS

This work was supported by the German Research Foundation (Grant No. Ri 1050/4-1). The authors thank Hansjörg Mast for his help in recording the MRI sequences. Author contributions were as follows: conceptualization, M.E., L.T., and F.B.; methodology, F.B. and M.B.; software, F.B. and M.B.; validation, F.B., M.E., L.T., and B.R.; formal analysis, F.B.; investigation, F.B., M.E., L.T., and M.B.; data curation, F.B.; writing original draft of the manuscript, F.B.; review and editing of the final manuscript, F.B., L.T., M.E., B.R., and M.B; visualization, F.B.; project administration, M.E. and B.R.; funding acquisition, M.E. and B.R.

AUTHOR DECLARATIONS

Conflict of Interest

The authors declare that there are no conflicts of interest to disclose regarding this work.

Ethics Approval

This study was approved by the Freiburg University Hospital Ethical Committee (No. 206/09). All subjects gave their written informed consent before inclusion in the study.

DATA AVAILABILITY

The data that support the findings of this study are available on request from the corresponding author, F.B.

[1]P. Laukka, P. Juslin, and R. Bresin, "A dimensional approach to vocal expression of emotion," Cogn. Emot. **19**, 633–653 (2005).
[2]P. H. Dejonckere, P. Bradley, P. Clemente, G. Cornut, L. Crevier-Buchman, G. Friedrich, P. Van De Heyning, M. Remacle, and V. Woisard, "A basic protocol for functional assessment of voice pathology,

especially for investigating the efficacy of (phonosurgical) treatments and evaluating new assessment techniques," Eur. Arch. Otorhinolaryngol. **258**, 77–82 (2001).
[3]I. R. Titze, R. Long, G. I. Shirley, E. Stathopoulos, L. O. Ramig, L. M. Carroll, and W. D. Riley, "*Messa di voce*: An investigation of the symmetry of crescendo and decrescendo in a singing exercise," J. Acoust. Soc. Am. **105**, 2933–2940 (1999).
[4]J. Sundberg, *Die Wissenschaft von der Singstimme (Forum Musikpädagogik)* [*The Science of the Singing Voice (Music Education Forum)*], 2nd ed. (Wißner, Augsburg, Germany, 2015).
[5]I. R. Titze, "On the relation between subglottal pressure and fundamental frequency in phonation," J. Acoust. Soc. Am. **85**, 901–906 (1989).
[6]B. Richter and M. Echternach, *Die Stimme: Grundlagen, Künstlerische Praxis, Gesunderhaltung* (*The Voice: Basics, Artistic Practice, Maintaining Health*), 2nd revised ed. (Henschel, Leipzig, Germany, 2014).
[7]I. R. Titze, *Principles of Voice Production* (Prentice-Hall, Englewood Cliffs, NJ, 1994).
[8]I. R. Titze, "Simulation of vocal loudness regulation with lung pressure, vocal fold adduction, and source-airway interaction," J. Voice **37**, 152–161 (2023).
[9]Z. Zhang, "Mechanics of human voice production and control," J. Acoust. Soc. Am. **140**, 2614–2635 (2016).
[10]G. Fant, *Acoustic Theory of Speech Production: With Calculations Based on X-Ray Studies of Russian Articulations* (Walter de Gruyter, Berlin, Germany, 1971).
[11]R. D. Kent, "Vocal tract acoustics," J. Voice **7**, 97–117 (1993).
[12]B. E. Lindblom and J. E. Sundberg, "Acoustical consequences of lip, tongue, jaw, and larynx movement," J. Acoust. Soc. Am. **50**, 1166–1179 (1971).
[13]M. Köberlein, P. Birkholz, M. Burdumy, B. Richter, F. Burk, L. Traser, and M. Echternach, "Investigation of resonance strategies of high pitch singing sopranos using dynamic three-dimensional magnetic resonance imaging," J. Acoust. Soc. Am. **150**, 4191–4202 (2021).
[14]E. Bresch and S. Narayanan, "Real-time magnetic resonance imaging investigation of resonance tuning in soprano singing," J. Acoust. Soc. Am. **128**, EL335–EL341 (2010).
[15]E. Joliveau, J. Smith, and J. Wolfe, "Vocal tract resonances in singing: The soprano voice," J. Acoust. Soc. Am. **116**, 2434–2439 (2004).
[16]J. Sundberg, "Articulatory configuration and pitch in a classically trained soprano singer," J. Voice **23**, 546–551 (2009).
[17]J. Sundberg, "Articulatory interpretation of the 'singing formant,'" J. Acoust. Soc. Am. **55**, 838–844 (1974).
[18]I. R. Titze, "Nonlinear source–filter coupling in phonation: Theory," J. Acoust. Soc. Am. **123**, 2733–2749 (2008).
[19]I. R. Titze, "Acoustic interpretation of resonant voice," J. Voice **15**, 519–528 (2001).
[20]T. Ikävalko, A.-M. Laukkanen, A. McAllister, R. Eklund, E. Lammentausta, M. Leppävuori, and M. T. Nieminen, "Three professional singers' vocal tract dimensions in operatic singing, kulning, and edge—A multiple case study examining loud singing," J. Voice (published online 2022).
[21]A. Mainka, I. Platzek, A. Klimova, W. Mattheus, M. Fleischer, and D. Mürbe, "Relationship between epilarynx tube shape and the radiated sound pressure level during phonation is gender specific," Logoped. Phoniatr. Vocol. **48**, 44–56 (2023).
[22]M. Leppävuori, E. Lammentausta, A. Peuna, M. K. Bode, J. Jokelainen, J. Ojala, and M. T. Nieminen, "Characterizing vocal tract dimensions in the vocal modes using magnetic resonance imaging," J. Voice **35**, 804.e27–804.e42 (2021).
[23]B. H. Story, I. R. Titze, and E. A. Hoffman, "The relationship of vocal tract shape to three voice qualities," J. Acoust. Soc. Am. **109**, 1651–1667 (2001).
[24]M. Echternach, F. Burk, M. Burdumy, L. Traser, and B. Richter, "Morphometric differences of vocal tract articulators in different loudness conditions in singing," PLoS One **11**, e0153792 (2016).
[25]M. Echternach, J. Sundberg, M. Markl, and B. Richter, "Professional opera tenors' vocal tract configurations in registers," Folia Phoniatr. Logop. **62**, 278–287 (2010).
[26]M. Echternach, J. Sundberg, S. Arndt, M. Markl, M. Schumacher, and B. Richter, "Vocal tract in female registers—A dynamic real-time MRI study," J. Voice **24**, 133–139 (2010).
[27]M. Echternach, J. Sundberg, T. Baumann, M. Markl, and B. Richter, "Vocal tract area functions and formant frequencies in opera tenors' modal and falsetto registers," J. Acoust. Soc. Am. **129**, 3955–3963 (2011).
[28]M. Bunch and J. Chapman, "Taxonomy of singers used as subjects in scientific research," J. Voice **14**, 363–369 (2000).
[29]M. Echternach, J. Sundberg, S. Arndt, T. Breyer, M. Markl, M. Schumacher, and B. Richter, "Vocal tract and register changes analysed by real-time MRI in male professional singers—A pilot study," Logoped. Phoniatr. Vocol. **33**, 67–73 (2008).
[30]M. Burdumy, L. Traser, B. Richter, M. Echternach, J. G. Korvink, J. Hennig, and M. Zaitsev, "Acceleration of MRI of the vocal tract provides additional insight into articulator modifications," Magn. Reson. Imaging **42**, 925–935 (2015).
[31]P. Boersma, "Praat, a system for doing phonetics by computer," Glot. Int. **5**, 341–345 (2001), available at https://www.researchgate.net/publication/208032992_PRAAT_a_system_for_doing_phonetics_by_computer#fullTextFileContent.
[32]F. Burk, "Untersuchung zur Vokaltraktmorphologie professioneller Sänger in verschiedenen Lautstärkekontexten vermittels dynamischer" (Echtzeit-Magnetresonanztomographie) ("Investigation of the vocal tract morphology of professional singers in different loudness contexts using dynamic real-time magnetic resonance imaging"), Doctoral dissertation, Albert–Ludwigs–Universität Freiburg im Breisgau, Freiburg, Germany, 2018.
[33]M. Spitzer, J. Wildenhain, J. Rappsilber, and M. Tyers, "BoxPlotR: A web tool for generation of box plots," Nat. Methods **11**, 121–122 (2014).
[34]S. Collyer, P. J. Davis, C. W. Thorpe, and J. Callaghan, "Sound pressure level and spectral balance linearity and symmetry in the *messa di voce* of female classical singers," J. Acoust. Soc. Am. **121**, 1728–1736 (2007).
[35]P. Birch, B. Gümoes, H. Stavad, S. Prytz, E. Björkner, and J. Sundberg, "Velum behavior in professional classic operatic singing," J. Voice **16**, 61–71 (2002).
[36]J. Sundberg, P. Birch, B. Gümoes, H. Stavad, S. Prytz, and A. Karle, "Experimental findings on the nasal tract resonator in singing," J. Voice **21**, 127–137 (2007).
[37]M. Havel, J. Sundberg, L. Traser, M. Burdumy, and M. Echternach, "Effects of nasalization on vocal tract response curve," J. Voice **37**, 339–347 (2023).
[38]M. Echternach, C. Högerle, M. Köberlein, P. Schlegel, M. Döllinger, B. Richter, and M.-A. Kainz, "The effect of nasalance on vocal fold oscillation patterns during the male passaggio," J. Voice **35**, 500.e9–500.e16 (2021).
[39]J. J. Jennings and D. P. Kuehn, "The effects of frequency range, vowel, dynamic loudness level, and gender on nasalance in amateur and classically trained singers," J. Voice **22**, 75–89 (2008).
[40]T. Shipp, "Vertical laryngeal position: Research findings and application for singers," J. Voice **1**, 217–219 (1987).
[41]M. Guzman, A. Lanas, C. Olavarria, M. J. Azocar, D. Muñoz, S. Madrid, S. Monsalve, F. Martinez, S. Vargas, P. Cortez, and R. M. Mayerhoff, "Laryngoscopic and spectral analysis of laryngeal and pharyngeal configuration in non-classical singing styles," J. Voice **29**, 130.e21–130.e28 (2015).
[42]C. T. Herbst, M. Hess, F. Müller, J. G. Švec, and J. Sundberg, "Glottal adduction and subglottal pressure in singing," J. Voice **29**, 391–402 (2015).
[43]A. Mainka, A. Poznyakovskiy, I. Platzek, M. Fleischer, J. Sundberg, and D. Mürbe, "Lower vocal tract morphologic adjustments are relevant for voice timbre in singing," PLoS One **10**, e0132241 (2015).
[44]J. Sundberg, R. Leanderson, and C. von Euler, "Activity relationship between diaphragm and cricothyroid muscles," J. Voice **3**, 225–232 (1989).
[45]J. Sundberg, R. Leanderson, C. von Euler, and E. Knutsson, "Influence of body posture and lung volume on subglottal pressure control during singing," J. Voice **5**, 283–291 (1991).
[46]L. Traser, A. C. Özen, F. Burk, M. Burdumy, M. Bock, B. Richter, and M. Echternach, "Respiratory dynamics in phonation and breathing—A real-time MRI study," Respir. Physiol. Neurobiol. **236**, 69–77 (2017).
[47]J. Iwarsson and J. Sundberg, "Effects of lung volume on vertical larynx position during phonation," J. Voice **12**, 159–165 (1998).
[48]M. Yadav, D. Cabrera, and D. T. Kenny, "Assessing linearity in the loudness envelope of the *messa di voce* singing exercise through acoustic signal analysis," J. Voice **29**, 646.e11–646.e21 (2015).