Characterizing Vocal Tract Dimensions in the Vocal Modes Using Magnetic Resonance Imaging

*,†Mari Leppävuori, †,‡,§Eveliina Lammentausta, †,‡Arttu Peuna, ‡Michaela K. Bode, ‖Jari Jokelainen, *,¶Juha Ojala, and †,‡,§Miika T. Nieminen, *†‡§‖*Oulu, and* ¶*Helsinki, Finland*

Summary: Objective. The aim was to study vocal tract dimensions in four vocal modes − Neutral, Curbing, Overdrive and Edge − from Complete Vocal Technique (CVT) by means of magnetic resonance imaging (MRI). Furthermore, the purpose was to test the feasibility of MRI to assess CVT vocal modes.

Methods. Four nonclassical singers (two females, two males) trained in CVT were imaged with an MRI scanner while singing sustained vowels at same pitch (Bb4 for females, F4 for males) in all vocal modes. Audio signals were simultaneously recorded through a pipe for quality assurance purposes. Auditory evaluation was performed by three CVT teachers in the scanner control room via headphones, and by one CVT teacher inside the MRI room. Previously developed measurement models modified by the authors were used to measure vocal tract dimensions from sagittal MRI projections. Repeatability test was performed for all measurements.

Results. In all subjects, vocal tract dimensions displayed differences between the vocal modes. Edge stood out from other vocal modes by showing most laryngeal narrowing accompanied by shortest vocal tract and highest vertical laryngeal position. For Neutral, least mouth opening and shortest distance between tongue and palate were found. Curbing differed consistently from Edge and somewhat from Overdrive showing higher measured values for vocal fold length. Differences regarding vocal fold length were also detected between Neutral and Edge. As expected, differences in vocal tract dimensions were found between samples sung with different vowels.

Conclusions. Vocal tract adjustments play a key role in the production of the vocal modes. The model used to measure vocal tract dimensions succeeded in finding significant differences between the vocal modes, also detecting differences between different vowel productions. The method used to characterize vocal tract dimensions seem promising and would be worthwhile to apply to a larger material.

Key Words: Vocal tract—Vocal folds—Vocal modes—Complete Vocal Technique—Magnetic resonance imaging.

INTRODUCTION

All voiced sounds are results of source-filter interaction with vibrating vocal folds providing a source impulse which is filtered by the vocal tract and radiated from mouth and nose.[1-5] Further, glottal flow and oscillatory characteristics of vocal folds are influenced by acoustic pressure of the vocal tract.[6-8] This complex interaction is regulated by articulators; lips, jaw, tongue, soft palate and larynx. Shaping the vocal tract results in changing the frequency response of the filter defining spectral energy distribution in the acoustic output. For human ear, this output signal is perceived as pitch, vowel, loudness and voice quality.[1-5,8-12]

Vowels are related to adjustments of tongue shaping the vocal tract. Two dimensions have been used to identify vowel spaces in the vocal tract; ‘open-close’ describing the space between tongue and palate, and ‘back-front’ describing the horizontal placement of tongue resulting in more or less pharyngeal space.[13] These dimensions determine the vowels which are traditionally classified on the basis of two lowest formant frequencies.[1-5,14] Besides tongue positioning, different vowels are also associated with differences in vertical laryngeal position, shape of the lower vocal tract as well as lip and jaw opening.[1,3] It has also been noticed that different vowels provoke different amounts of exhaled air during phonation influencing subglottal pressure, opening of the glottis and condition of the vocal folds.[3] Moreover, variation in the length, tension and elasticity of the vocal folds have been related to control of pitch, loudness and tone quality.[1-3,12,15-18]

Remarks have been made on how singers modify vowels to produce particular sounds,[4,12] how some vowels work better with certain pitches, singings styles and voice qualities,[7,19] and how different singing styles have been found to benefit from differently shaped vocal tracts. For example, ‘belting’, as a loud and brassy sound, has shown megaphone shape with narrow pharynx and wide mouth opening, which have been found to be best accomplished by using vowels [e:] or [æ:].[9,20] As the opposite, inverted megaphone-shaped vocal tract with lip-rounding vowels and stretched pharyngeal space have been described for sounds used in classical singing[9] and in ‘head voice’.[20] Moreover, it has been reported that vowels with laryngeal narrowing, like [e:], [æ:], are apt to be pronounced in ‘sharp voice’, whereas ‘soft voice’ has been associated with vowels like [u:] and [i:] that provide more

Accepted for publication January 16, 2020.
Declarations of interest: None.

From the *Community of Research in Education, Music, and the Arts, Faculty of Education, University of Oulu, Oulu, Finland; †Research Unit of Medical Imaging, Physics and Technology, Faculty of Medicine, University of Oulu, Oulu, Finland; ‡Department of Diagnostic Radiology, Oulu University Hospital and University of Oulu, Oulu, Finland; §Medical Research Center, Oulu University Hospital and University of Oulu, Oulu, Finland; ‖Infrastructure for Population Studies, Faculty of Medicine, University of Oulu and Unit of Primary Health Care, Oulu University Hospital, Oulu, Finland; and the ¶DocMus Doctoral School, Sibelius Academy, University of the Arts Helsinki, Helsinki, Finland.
Address correspondence and reprint requests to Mari Leppävuori, Community of Research in Education, Music, and the Arts, Faculty of Education, University of Oulu, P.O. Box 8000, Oulu FI-90014, Finland
E-mail: mari.leppavuori@student.oulu.fi

Journal of Voice, Vol. ■■, No. ■■, pp. ■■−■■
0892-1997

https://doi.org/10.1016/j.jvoice.2020.01.015

laryngeal space. However, vowels can also be produced with different tone qualities and phonation types in different registers, which in turn constitute a factor of the vowel quality.[3] Yet, in a study of seven male singers singing same pitches and vowels in their chest and falsetto registers, no evidence of a uniform resonance tuning strategy among the singers was found.[21]

Loud sounds have been associated with an increase in high-frequency energy.[22-24] which has been reported to correlate with wider hypopharyngeal area and lower vertical laryngeal position for classically trained singers.[25-27] Correspondingly, higher degrees of laryngeal and pharyngeal narrowing,[20,24] and both higher[20,27,28] and lower[24,28] vertical laryngeal position have been found for loud nonclassical singing. Furthermore, interindividual differences in vocal tract adjustments for loud sounds in nonclassical singing styles have been reported,[28] some of which might be related to individual morphology and articulatory maneuvers,[26,29] as well as to different subcategories of belting,[30-33] or to metallic voice quality, which similarly to loud sounds, is associated with an increase in high-frequency energy[22,34] accompanied by laryngeal rise, and laryngeal and pharyngeal narrowing.[22]

Complete Vocal Technique (CVT) is one of the methods used in voice pedagogy, and just as with other teaching concepts, CVT terminology and related activities should be critically assessed to test the assertions, bases, structures and practices of the method and to facilitate and promote common understanding and learning.[35-37] CVT divides all the voiced sounds (with vowels) into 'non-metallic', 'reduced metallic' and 'full metallic' sounds and further into four 'vocal modes' called 'Neutral', 'Curbing', 'Overdrive' and 'Edge'. According to CVT, any voiced sound (with a vowel) produced by human voice can be analysed as one of the vocal modes. Moreover, each of the vocal modes are identified, characterized and defined according to auditory attributes – 'metal', 'density', vowel and loudness. In general, all vocal modes are used in all genres of music.[30,38-43]

According to CVT, 'a distinct metallic tone could be described as a harder, more raw or direct sound' whereas 'density' refers to the 'degree of "compactness", the degree to which the note is "filled out", how "solid" or "weighty" it is, or perhaps how much "core" or "foundation" the note has'.[38] Neutral is the nonmetallic vocal mode and sounds softer or milder compared to the metallic modes. Neutral is used when singing or speaking quietly, but it can also be loud in higher pitches. Curbing is the reduced metallic vocal mode with a restrained character, and it is used when singing or speaking with a slightly plaintive and held-back voice with a relatively loud volume. Compared to Curbing, Overdrive and Edge can be louder, ranging from reduced to full metallic. The character of Overdrive can be described as shouty, direct and unrestrained, whereas Edge is sharp, clear and bright. Regarding pitch, Neutral is used in all parts of the voice, Curbing and Edge are used up to C6 (1046 Hz) for females and in all parts for males and Overdrive up to D5 (587 Hz) for females, and C5 (523 Hz) for males, respectively. Furthermore, Neutral, Overdrive and Edge can be altered regarding density resulting in either 'fuller' or 'reduced density', whereas Curbing is defined as a reduced density vocal mode[38] (Table 1).

In CVT pedagogy, vowels and vowel variations play an important role in the practical use of the vocal modes, regulating loudness as well as the degree of metal and density in the sound. The metallic vocal modes – Curbing, Overdrive and Edge – are restricted to certain vowels. According to CVT, vowels [u:], [o:] and [ʌ:] (as in 'hungry') belong to Curbing, vowels [ε:] and [ɔ:] (as in beginning of 'so') to Overdrive, and vowels [ɪ:], [ε:], [æ:] and [ú:] to Edge. In nonmetallic Neutral all vowels can be used. However, vowels [i:], [u:] and [ɑ:] are recommended when practicing nonmetallic sounds.[38]

In previous research of CVT vocal modes, laryngostroboscopic studies have demonstrated visible differences regarding laryngeal gestures between Overdrive, Edge and Curbing, with corresponding differences in long-term average spectrum, electroglottography and acoustic measures.[30,41] Edge was found to have most constricted laryngeal setting, whereas Overdrive showed less and Curbing least laryngeal constriction exposing more of the vocal folds both anterior-posteriorly and laterally. Vertical laryngeal position was found to be highest in Edge, and lowest in Overdrive. In acoustic analysis, Edge has shown more energy from around 2300 Hz to around 3500 Hz compared to Overdrive.[30] For Curbing, less energy in higher frequencies starting from 3 kHz was detected in comparison to Overdrive and Edge,[41] whereas least high-frequency energy was found for Neutral when compared to other vocal modes.[39] Furthermore, Overdrive has shown higher degree of subglottal pressure and vocal fold adduction when compared to tones sung in falsetto. It was also suggested that Overdrive is produced with firmer contraction of vocalis muscle resulting in thicker vocal folds and higher contact quotient compared to falsetto. Overdrive has also been found to have higher first formant tuned to second harmonic.[44] Finally, reduced density has been associated with laryngeal tilt and decrease of higher spectral energy.[43] As far as the authors of this paper are aware, no studies have been made to characterize vocal tract dimensions in the four vocal modes by quantitative research methods. Further research regarding vocal modes,[30,41] metallic voice quality,[22] density[43] and clarification on how different loudness conditions influence vocal tract in nonclassical singers[26] have been called for.

Magnetic resonance imaging (MRI) has offered valuable information on vocal tract dimensions and articulators in relation to different singing styles,[20,45] timbre,[25] registers,[19,46-49] voice qualities,[50] vowels,[51] pitch changes,[49] resonance[52] and loudness,[26,49] and also a safe way for the singers to be examined without exposing them to ionizing radiation. MRI mid-sagittal projections have been used to depict vocal 0.02w?>tract anatomy and anterior-posterior dimensions of the cavities, see, for example.[26,48,53,54]

This study examines vocal tract dimensions in the cases of two female and two male nonclassical singers, all trained in

TABLE 1.
Overview of the Vocal Modes Regarding Metal, Density, Loudness and Character, and Vowel Selection (Used in This Study) According to CVT Description and International Phonetic Alphabet (IPA) Correlates.[13] Singers Used Various Vowels in Curbing, Denoted Here With Superscripts. All Samples Were Labelled With First Letter of Vocal Mode Name (N, C, O, E) and Sample Number (1, 2), Later to Be Used for Image and Audio Recording Identification, See Figure 2 and Supplementary Material, Audio Recordings

Vocal Mode	Metal	Density	Loudness	Character	Selected Vowels	Phonetics IPA	Labels
Neutral	Nonmetallic	From reduced to full density	From very quiet to medium quiet*	Soft	'Ee' as in 'see'	close front vowel [iː]	N1
					'Oo' as in 'you'	close back vowel [uː]	N2
Curbing	Reduced metallic	Reduced density	From medium quiet to medium loud	Restrained	'Uh' as in 'hungry' [F1, F2, M1]	open-mid back vowel [ʌː]	C1
					'Eh' as in 'stay' [M2]	open-mid front vowel [ɛː]	C1
					'Oh' as in the beginning of 'so' [F1]	open-mid back vowel [ɔː]	C2
					'O' as in 'woman' [F2, M2]	close-mid back vowel [oː]	C2
					'I' as in 'sit' [M1]	near-close near-front vowel [ɪː]	C2
Overdrive	From reduced to full metallic	From reduced to full density	From medium loud to very loud	Clear/shout	'Eh' as in 'stay'	open-mid front vowel [ɛː]	O1
					'Oh' as in the beginning of 'so'	open-mid back vowel [ɔː]	O2
Edge	From reduced to full metallic	From reduced to full density	From medium loud to very loud	Clear/sharp	'Eh' as in 'stay'	open-mid front vowel [ɛː]	E1
					'A' as in 'and'	near-open front vowel [æː]	E2

* Can be loud in higher pitches.

CVT, by means of MRI sagittal projections. The specific aims of the study were to investigate (1) how vocal tract configurations differ between the vocal modes, and (2) what kind of trends, if any, these differences show between singers.

MATERIAL AND METHODS

Thirty-two MRI sagittal projections, 8 from each singer and 2 from each vocal mode were examined with 24 measurements, of which 15 were used to characterize vocal tract dimensions in 4 vocal modes. The remaining nine measurements served as auxiliary lines and frames allowing the vocal tract measurements to be carried out by a segmentation tool. All measurements were used to test repeatability.

Ethical issues

The Regional Ethics Committee of the Northern Ostrobothnia Hospital District approved the study (decision 49/2015 § 135). All subjects gave written consent and they had the possibility to interrupt their participation in the study at any stage. The method used in this research did not cause any inconveniences or risks to examinees.

Participants

Two female (F1 and F2) and two male (M1 and M2) authorized CVT teachers and singers volunteered as subjects. They had 7-10 years of experience in using CVT method in teaching and in their own artistic training. The age of subjects ranged from 33 to 51 years and they all represented nonclassical, western popular styles of singing.

All participants were examined by phoniatrician for characterization of vocal tract anatomy. Deviating from other singers, laryngeal asymmetry was observed for M2; larynx and vocal folds were anterior-posteriorly skewed. Even though this asymmetry was considered as normal, it was taken into account when processing the data.

Vocal task

The singers were instructed to give sound examples in all four vocal modes on Bb4 (466 Hz) for females and F4 (349 Hz) for males. The chosen pitches represent higher part of the voice where all the vocal modes can be used. Vowels and loudness (Table 1) for the vocal mode samples were selected according to recommendations by CVT and where the subjects felt it was comfortable to produce the vocal modes. Compared to other vocal modes, more vowels were used in Curbing due to different individual experiences in producing the vocal mode in a stable manner. The order for performing the vocal samples was chosen by the participant. Regarding duration and timing of each sample, the singers were instructed to start the sound 1 second before the scanning started and sustain it for approximately 14 seconds until the scanning was finished.

Imaging procedure

All participants were examined using Siemens Skyra 3T MRI device (Siemens Healthcare, Erlangen, Germany) with a 24-channel head/neck coil and 18-channel body matrix coil. Volunteers were positioned in supine position with ear plugs and headphones on. The head was supported by foam wedges to minimize lateral movement. The body matrix coil was positioned on top of the chest as close to larynx as possible. 3D VIBE CAIPIRINHA sequence was used (TR = 4.21 ms, TE = 2.58 ms, Field of View 450 × 337 mm, matrix size 320 × 216 and slice thickness of 1.4 mm, yielding, isotropic resolution of 1.4 mm, scan time of 13 seconds). Due to extensive imaging protocol, singers were allowed to have short breaks, followed by repositioning and new localizer scans.

Audio recording and evaluation procedure

The entire imaging session was recorded using a plastic hose (length 5 m, inner diameter 15 mm) attached to the head/neck coil, approximately 3 cm from the participant's mouth. Foam plastic pieces were inserted into both ends of the hose to attenuate flutter echoes generated within the hose. Outside the scanner room, the other end was attached to a Behringer ECM8000 condenser microphone (Behringer, Germany), Tascam US-122 audio interface (TEAC Corporation, Japan) and a laptop computer with Audacity software (The Audacity Team). The sounds were recorded in 32-bit samples with sampling rate of 44.1 kHz and saved in noncompressed.WAV files. The recording levels were adjusted for each singer to gain maximal range of dynamics while avoiding clipping and distortion (Supplementary Material, Research devices and procedures, Illustration 1a). For scanning and audio recording, each pitch was played to singers through headphones (intercom) as a cue to start singing. To avoid motion artefacts in the images, scanning started only after the onset of each vocal sample (Supplementary Material, Research devices and procedures, Illustration 1b).

Three different approaches were used in real time evaluation procedure: (1) singer's own somatosensory feedback of the vocal mode performance, (2) auditory evaluation performed by one monitor (CVT teacher) inside the MRI room and (3) auditory evaluation performed by three monitors (CVT teachers) in the MRI scanner control room via audio recording set up (Supplementary Material, Research devices and procedures, Illustration 1). The singers were asked to practice the requested samples under the scanning conditions (with and without scanning noise) and to share their sensations after each vocal performance. Based on earlier studies on effects of singer education on auditory and kinaesthetic feedback[55] and neural control of singing,[56] this feedback was considered the main criterion in the evaluation procedure when approving the recorded samples for further analysis. Vocal mode production was further controlled by consensus of four other CVT trained singers and teachers. One CVT teacher assessed the vocal samples inside the MRI room with and without earplugs during the absence of MRI noise, and with earplugs during MRI noise, being able to compare each vocal mode sample in both evaluation conditions. Three CVT teachers in the MRI control room were able to evaluate the vocal samples with and without MRI noise through the recording set up (including plastic hose) and quality stereo headphones (AKG K121 Studio, AKG K240 DF and AKG K44, AKG Acoustics by Harman International Industries, United States). In this live monitoring procedure, each vocal mode sample was recorded 1-4 times until consensus was reached between all monitors and the singer.

The frequency response of the plastic hose (used in the recording set up) was measured using bursting balloons as impulse sources[57] (since devices producing frequency sweep with sufficient accuracy cannot be used in MRI environment). According to the results, high frequencies are attenuated in the hose, considerably. Power spectrum displays attenuation of approximately 10 dB per octave (Figure 1). It should be considered, however, that the nonflat impulse

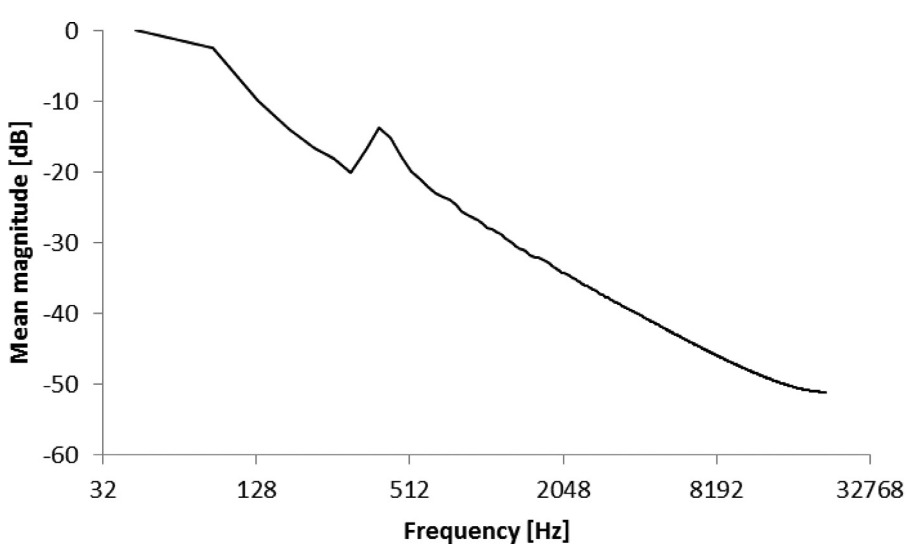

FIGURE 1. Frequency response of the plastic hose.

power spectrum of balloons displays less power at high frequencies.[57]

Since three monitors in the MRI scanner control room were exposed to the interference of the plastic hose attenuating the audio spectrum considerably, especially regarding higher frequency region, a further auditory evaluation was carried out to test the recognizability of the vocal modes. All the audio samples used in this study were evaluated by seven authorized CVT teachers. The order of the audio samples was randomized in the test and the participants performed the test independently with sound reinforcement system selected and adjusted as per their own choice and preferences (Supplementary Material, Research devices and procedures), and thus, with an unlimited number of repeats for each audio sample. According to the test results, the average recognizability percentage was 82%, showing highest value for Neutral with 98%, and lowest for Edge with 61%. Between singers, the samples from Male 2 displayed highest recognizability percentage with 89%, and Male 1 the lowest with 75%. Listener-specific percentages ranged between 59% and 97% (Supplementary Material, Auditory evaluation results).

Given that Edge has shown more energy in higher frequencies compared to other vocal modes,[30,39,41] and that the recording set up with the plastic hose interferes mostly with higher frequency region (Figure 1), it makes sense that according to further auditory evaluation results, Edge displayed lowest recognizability percentage in comparison to other vocal modes. Despite the apparent deterioration of the audio signal due to MRI noise, earplugs and the recording procedure involving a plastic hose, the singers were able to complete the tasks approved by all monitors and singers. Furthermore, considering that three approaches were used to assess vocal mode samples and that singer's own somatosensory feedback was considered the main criterion, the real-time evaluation condition for three monitors via recording set up seems sufficient.

From each vocal mode, two samples with different vowels were chosen for the image analysis to examine vocal tract adjustments in the four vocal modes (Figure 2, Table 1, Supplementary Material, Audio Recordings). These samples were selected based on the real time evaluation and thus, approved by all three evaluation approaches.

Image analysis

MRI sagittal projections were analysed using a modified template based on previous studies.[15,26,48,53,54,58] A sagittal projection with thickness of 3 mm through the midsection of the subject's head and neck was manually generated from the 3D image stack. Anatomical structures − septum, spina, teeth, uvula, spine, tongue, jawbone, epiglottis, arytenoid cartilages and vocal folds − were used as landmarks in detecting the sagittal centre line of the vocal tract. Outline of the vocal tract and 10 anatomical landmarks were manually drawn into each sagittal projection (Figure 3a, Table 2). Regarding anterior and posterior ends of vocal folds, landmarks were drawn to points which were clearly distinguishable from other soft tissue. Anatomical accuracy of manual segmentations was verified by a highly experienced neuroradiologist (MKB) in a blinded manner. Based on manual segmentation, auxiliary lines and measurements for vocal tract dimensions (distances and angles) were conducted by using an in-house MATLAB application (Mathworks, Natick, MA, USA) (Figure 3b, Table 2).

Repeatability

Repeatability test was carried out in two working phases: (1) by repeating the manual work of generating sagittal projection from the same 3D image stack three times, followed by the measurement procedure for each projection and (2) by repeating the measurement procedure three times on the same sagittal projection. For the manual segmentation work regarding one of the participants (M2), adjustments were made because of an asymmetric vocal tract structure. It was noticed that when the focus in generating the sagittal projection was kept in the centre of the laryngeal structures (vocal folds, arytenoid cartilages, epiglottis), the middle points of the upper vocal tract structures (septum, spina, teeth, uvula, spine, tongue, jawbone) did not settle in the same centreline. Due to this, adjustments for the sagittal projections were carried out as follows; as a starting point, the centreline of the laryngeal structures and septum was selected, and the rest of the vocal tract selection was adjusted accordingly. However, because of the clear asymmetry, a test was executed to compare sagittal projections generated using different approaches. The same 3D image stack was used for this comparison test; one projection was generated by selecting the centreline of the laryngeal structures and another projection was generated by selecting the centreline of the upper vocal tract structures. These two projections were then compared to the original sagittal projection with the above-mentioned adjustments.

Repeatability was evaluated using three thresholds. Good repeatability was determined by the coefficient of variation percentage (CV%) <10, moderate repeatability between 10 and 20 and poor repeatability >20.[59,60]

Intermodal, intramodal and interindividual comparison

Intermodal comparison was performed using statistical analysis to investigate the differences between the measured vocal tract dimensions of the vocal modes, and to see whether these differences show any trends among the singers. For this statistical comparison, repeated measures and compound symmetry matrixes were used. The analysis was done on 13 measured variables with good and moderate repeatability. In this comparison, vowels were disregarded; all samples representing same vocal mode were treated as one vocal mode category. Statistical analysis repeated measures were conducted using SAS procedure MIXED (version 9.4 TS Level 1M5, SAS Institute, Cary, NC). P values <0.01

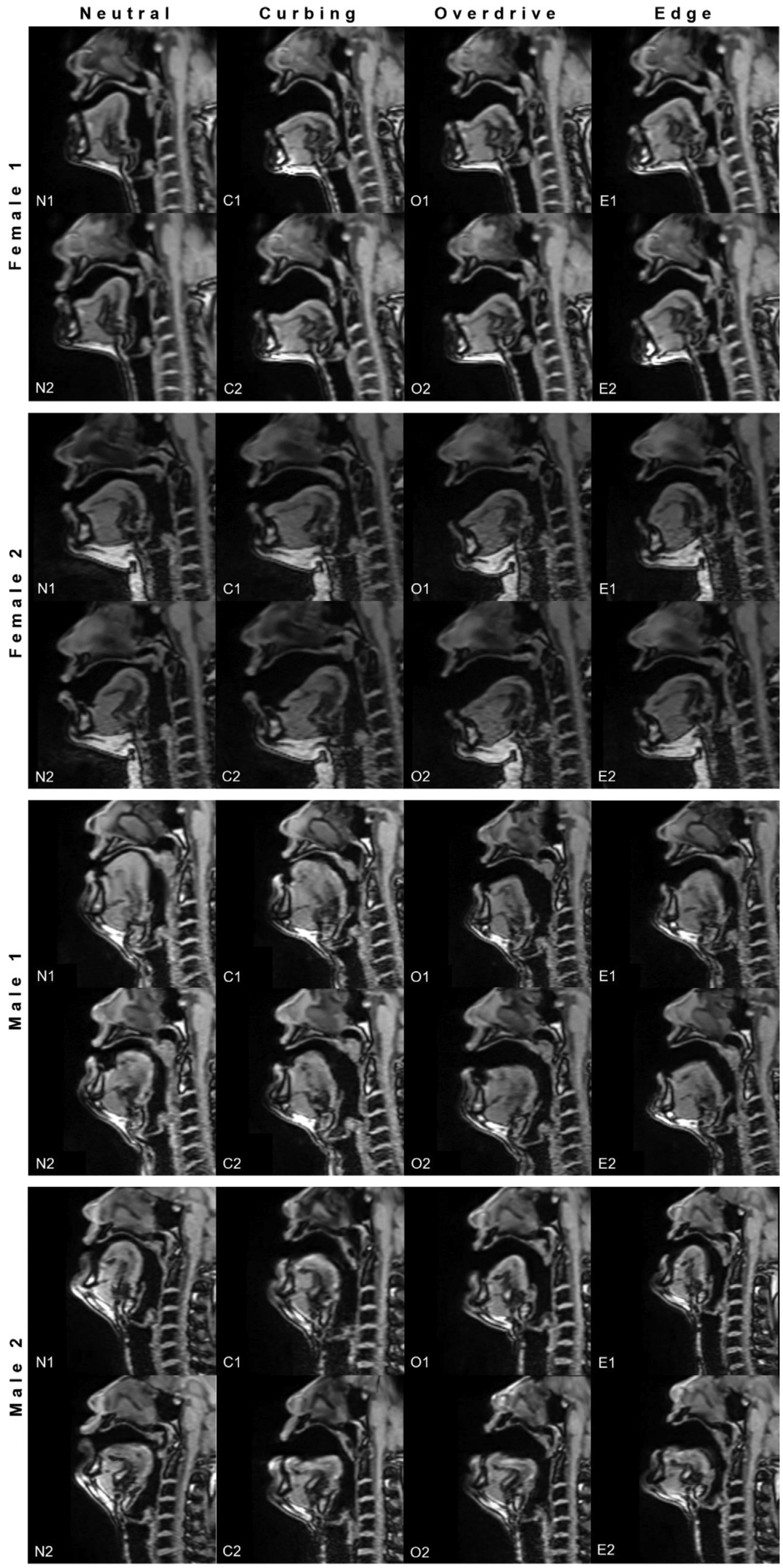

FIGURE 2. Singer profiles. MRI mid-sagittal projections of vocal mode samples from all singers; Female 1, Female 2, Male 1 and Male 2. Each image has been labelled with first letter of vocal mode name (N, C, O, E) and sample number (1, 2). For vowel descriptions, see Table 1.

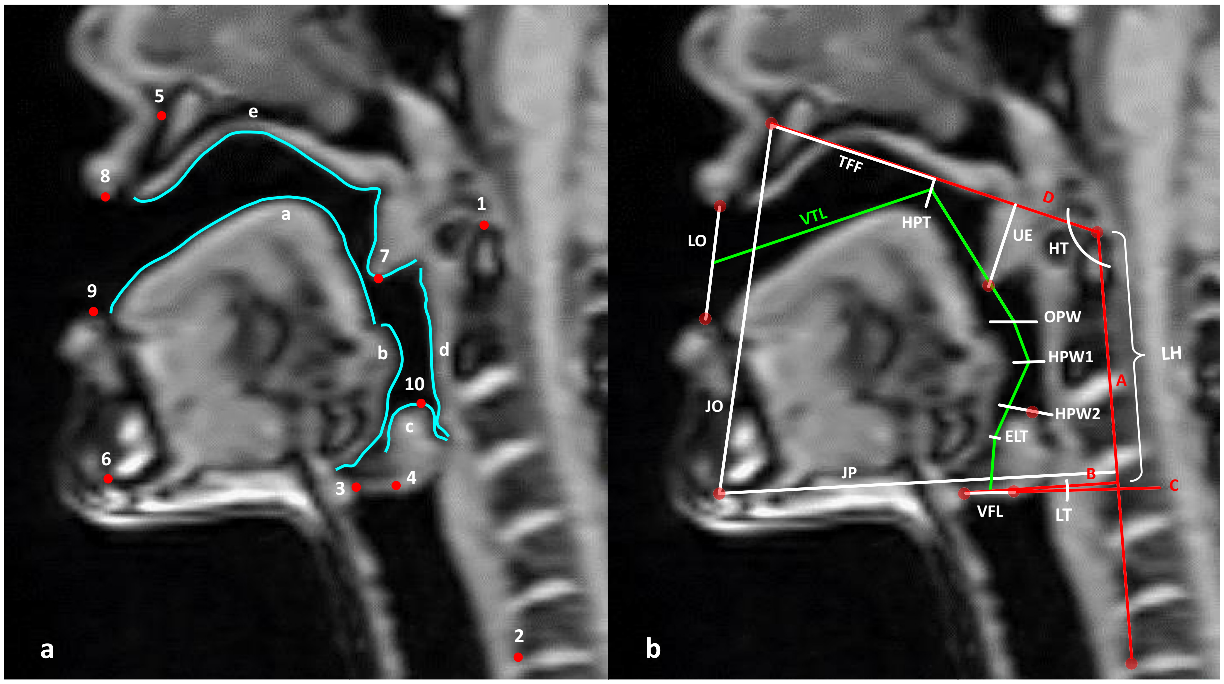

FIGURE 3. Segmentation process in two stages: a. manual segmentation, b. generated features based on manual segmentation (Table 2).

were considered statistically significant and <0.05 almost significant.

Intramodal comparison was performed to investigate the differences of the measured vocal tract dimensions between two vowels representing the same vocal mode. Absolute measurement results were compared as changes in percentages. Curbing was left out from the analysis due to inconsistency in vowel selection.

For interindividual comparison, singer-specific review was done to examine individual vocal tract physiology by (1) averaging the measured values for vocal tract length (VTL), larynx height (LH) and vocal fold length (VFL), and (2) comparing the absolute measured values to investigate the individual adjustments regarding vocal modes and vowels.

RESULTS

Repeatability

Twenty-four measurements were used to test the repeatability of the manual segmentation process (Table 3). Most variables displayed good repeatability in all comparisons. Poor repeatability was shown in two vocal tract measurements, epilaryngeal tube (ELT) and laryngeal tilt (LT). Image resolution of 1.4 mm was insufficient for these measurements and they were left out from further analysis. Moreover, measurements for auxiliary lines and angles were only used for the repeatability test and were not part of further analysis regarding vocal tract dimensions in the vocal modes.

Intermodal differences

Intermodal differences were found in 10 measurements (Figure 4). Neutral stood out from other vocal modes with lowest measured values in the distances between upper and lower lip (LO), upper and lower jaw (JO) and tongue and palate (HPT). Furthermore, Neutral differed from Overdrive and Edge by showing higher measured values in the shortest distance between epiglottis and pharyngeal wall (HPW1). Edge was distinguished from other vocal modes by most anterior-posterior narrowing just above the arytenoids (HPW2) accompanied by highest vertical laryngeal position (LH) and shortest vocal tract (VTL). When compared to Neutral and Curbing, Edge showed higher measured values for lip opening (LO), and lower measured values in the shortest distance between epiglottis and pharyngeal wall (HPW1) as well as for VFL. Curbing differed from Edge and Overdrive with higher measured values in VFL. Curbing also displayed somewhat higher measured values in uvula elevation (UE) in comparison to Overdrive. For head tilt (HT), lowest measured values were found in Neutral and highest in Edge. Measured values for jaw protrusion (JP), tongue from front (TFF) and oropharynx width (OPW) varied with no significant differences between the vocal modes.

For all singers, vowel [ɛ:] was used in two vocal modes, Overdrive and Edge. Compared to Overdrive, less space between tongue and palate (HPT), and between epiglottis and pharyngeal wall just above the arytenoids (HPW2) was shown for all singers in Edge. Furthermore, LH displayed lower measured values and head tilt (HT) higher measured values in Edge compared to Overdrive (Figure 5).

Intramodal differences

Intramodal differences were found in all vocal modes and in all 13 measured variables with good or moderate

TABLE 2.
Labels for Segmentation Features. See Figure 3

Labels for Manually Segmented Features (Figure 2a)				Labels for Features Generated by an in-House MATLAB Application (Figure 2b)			
Dots		Lines		Auxiliary Lines		Distances and Angles	
1	Cranial-most part of dens axis C2	**a**	Tongue	**A**	Caudo-anterior edge of C6 – cranial-most part of dens axis C2	**LO** lip opening	Vertical distance between lowest point of upper lip and highest point of lower lip
2	Caudo-anterior edge of C6	**b**	Epiglottis	**B**	Anterior commissure perpendicular to auxiliary line A	**JO** jaw opening	Distance between anterior end of hard palate and lowermost edge of jawbone contour
3	Anterior commissure	**c**	Soft tissue around arytenoid cartilages	**C**	Anterior commissure – posterior ends of the vocal folds	**HPT** highest point of tongue	Shortest distance between tongue and palate according to a line perpendicular to auxiliary line D
4	Posterior ends of vocal folds	**d**	Mucosal cover of spine	**D**	Cranial-most part of dens axis (C2) – anterior end of the hard palate	**UE** uvula elevation	Distance between auxiliary line D and lowermost part of uvula contour according to a line perpendicular to auxiliary line D
5	Anterior end of hard palate	**e**	Palate			**OPW** oropharynx width	Shortest dorsal-ventral distance between posterior contour of tongue and mucosal cover of spine
6	Lowermost edge of jawbone contour					**HPW1** hypopharynx width 1	Shortest distance between epiglottis and mucosal cover of spine
7	Lowermost part of uvula contour					**HPW2** hypopharynx width 2	Distance between epiglottis and mucosal cover of spine above the highest point of soft tissue around arytenoid cartilages
8	Lowest point of upper lip					**ELT** epilaryngeal tube	Shortest distance between epiglottis and soft tissue around arytenoid cartilages
9	Highest point of lower lip					**VFL** vocal fold length	Distance between anterior commissure and posterior end of vocal folds (clearly distinguished from other soft tissue)
10	Highest point of soft tissue around arytenoid cartilages					**VTL** vocal tract length	Measured as five separate distances using the middle points of LO, HPT, OPW, HPW1, ELT and VFL measurement lines

(*Continued*)

TABLE 2. (*Continued*)

Labels for Manually Segmented Features (Figure 2a)		Labels for Features Generated by an in-House MATLAB Application (Figure 2b)		
Dots	Lines	Auxiliary Lines	Distances and Angles	
			LH larynx height	Distance from cranial-most part of dens axis to the point where auxiliary lines A and B cross
			JP jaw protrusion	Distance between lowermost edge of jawbone and perpendicular intersection point relative to auxiliary line A
			TFF tongue from front	Distance between anterior end of hard palate and intersection point of auxiliary line D and HPT
			LT laryngeal tilt	Angle between auxiliary lines B and C
			HT head tilt	Angle between auxiliary lines D and A

repeatability. Among the singers, similar changes from one vowel to another within Neutral, Overdrive and Edge were found in eight measured variables related to mouth opening (LO, JO), tongue position (OPW, TFF), VTL, LH and pharyngeal narrowing (HPW1, HPW2) (Figure 6). In Neutral, less mouth opening (LO, JO) and a more posterior tongue position (OPW, TFF) were found in [u:] compared to [i:]. For Overdrive [ɔ:], longer vocal tract (VTL), more jaw opening (JO), and more space between epiglottis and pharyngeal wall just above the laryngeal inlet (HPW2) were shown in comparison to [ɛ:]. In Edge, shorter distance between epiglottis and pharyngeal wall (HPW1), higher vertical laryngeal position (LH) and more mouth opening (LO, JO) were detected in [æ:] compared to [ɛ:]. The shortest distance between tongue and palate was found to be further from the anterior end of hard palate (TFF) in Overdrive [ɔ:] compared to [ɛ:], and in Edge [æ:] compared to [ɛ:], respectively. The magnitude of these changes varied among the singers.

Interindividual differences

In the singer-specific review, the average values for VTL, LH and VFL displayed differences between the singers (Figure 7). F1 was found to show the lowest values followed by F2, M2 and finally M1 with the highest values. For the absolute measured values of these variables, some interindividual differences were found between the vocal modes and vowel specific adjustments (Figure 8). For F1, longest vocal tract (VTL) was detected in Neutral and only little variation between Curbing, Overdrive and Edge. Similarly, M1 showed moderate variation with none of the vocal modes standing out. For F2 and M2, greater intermodal and intramodal differences were found in VTL. Regarding laryngeal height (LH), F1, F2 and M2 showed lowest measured values in both Edge samples compared to other vocal modes, whereas for M1 the difference was not that clear. Lowest vertical laryngeal position was found in Neutral for F1, in Curbing for F2 and M2, and in Overdrive for M1. Finally, F1 was shown to gradually decrease VFL from Neutral to Edge, while for other singers the highest values were measured in Curbing.

DISCUSSION

This study is the first attempt to examine vocal tract dimensions in the four vocal modes from CVT by means of MRI. As a result, some significant differences in the measured vocal tract dimensions between the vocal modes were found. Neutral and Edge stood out from other vocal modes in several comparisons, while Curbing and Overdrive showed some differences, mainly when compared to one of the vocal modes. As expected, differences in vocal tract dimensions were found between samples sung with different vowels. Based on the auditory evaluation, interindividual differences were detected, some of which might be reflected in the results.

TABLE 3.
Repeatability of Generating Sagittal Projection and Segmentation Work Presented as Coefficient of Variation Percentages (CV%). Regarding Asymmetric Vocal Tract Structure (M2), Different Strategies in Generating Sagittal Projection (Described in Repeatability) Were Also Compared

		Sagittal Projection			Segmentation		Sagittal Projection, Asymmetric Vocal Tract
	MRI Measurements	F1	F2	M2	F1	F2	M2
Vocal tract dimensions	LO Lip opening	1.60	0.83	2.99	1.73	0.32	8.30
	JO Jaw opening	0.15	0.34	0.88	0.06	0.22	0.39
	HPT Highest point of tongue	2.45	1.26	2.17	4.42	0.00	1.59
	OPW Oropharyngeal width	9.78	4.36	1.89	8.22	2.72	4.42
	HPW1 Hypopharyngeal width 1	11.23	3.18	0.64	10.30	1.86	6.05
	HPW2 Hypopharyngeal width 2	4.07	3.66	3.09	3.71	1.46	3.57
	ELT Epilaryngeal tube	35.36	17.16	0.00	35.36	3.29	33.69
	VFL Vocal fold length	5.57	3.77	2.71	5.59	4.74	7.09
	VTL Vocal tract length	0.17	0.26	0.43	0.13	0.31	0.98
	LH Larynx height	0.00	0.23	1.58	0.13	0.06	1.22
	JP Jaw protrusion	0.05	0.20	0.44	0.11	0.11	0.65
	TFF Tongue from front	1.32	2.31	1.75	1.15	0.45	1.24
	LT Laryngeal tilt	−668.73	−18.97	−37.04	−668.73	−2.58	−162.75
	UE Uvula elevation	1.34	0.39	6.56	0.95	0.00	4.60
	HT Head tilt	0.07	0.34	0.12	0.07	0.09	1.06
Auxiliary lines and angles	Length of line a	1.27	0.17	0.96	0.79	0.14	1.12
	Length of line b	1.27	1.90	5.66	0.91	0.93	3.67
	Length of line c	9.48	6.79	18.10	3.24	0.77	16.84
	Length of line d	1.93	0.31	9.10	3.14	0.09	1.30
	Length of line e	0.75	0.13	0.88	0.70	0.23	0.58
	Length of auxiliary line A (from cranial-most part of dens axis C2 to the intersection point of JP)	0.33	0.34	1.45	0.29	0.27	2.62
	Length of auxiliary line D	0.58	0.27	2.09	0.26	0.09	0.27
	Angle between JO and JP	0.37	0.15	1.51	0.22	0.05	1.00
	Angle between JO and auxiliary line D	0.26	0.29	1.44	0.26	0.06	0.98

Intermodal differences

Edge showed most pharyngeal anterior-posterior narrowing accompanied by highest vertical laryngeal position compared to other vocal modes, while Neutral, Curbing and Overdrive did not settle in any particular order regarding these measurements. Our findings are in line with previous laryngostroboscopic observations of Edge but contradict with the results from the same studies reporting that Overdrive was found to display lowest vertical laryngeal position compared to Curbing and Edge, while Curbing was found to show least laryngeal narrowing compared to Overdrive and Edge.[30,41] Many studies have reported similar findings on relatively high larynx with laryngeal narrowing related to belting,[2,8,20,24,27,28,61] sharp voice[3] and metallic voice.[22] However, vertical laryngeal position has also been found to be lower in loud and higher in soft phonation for nonclassical singing styles.[24] Similarly, in comparison to 'mix' voice production, both laryngeal lowering and unchanged laryngeal height has been reported for belting production.[28] These findings, in addition to our results, suggest that LH can be altered within loud sound production, presumably causing changes in other auditory attributes, such as vowel characteristics.[25]

Neutral was found to represent the inverted megaphone-shaped vocal tract with least mouth opening (LO, JO) and most oral narrowing (HPT) compared to other vocal modes. Regarding pharyngeal and laryngeal narrowing, megaphone-shaped and also shortest vocal tract, when compared to other vocal modes, was most clearly represented in Edge, where vowels [e:] and [æ:] were used. These findings on the differences between Neutral and other vocal modes correspond to descriptions on other phonation types. Shorter vocal tract with widened mouth opening has been previously observed for twangy sound when compared to normal and yawny voice qualities.[50] Wider mouth opening has also been related to megaphone-shaped vocal tract and open front vowels, such as [e:] and [æ:],[9,20] which have been noticed to provoke laryngeal narrowing resulting in 'sharp voice'.[3] In contrast, inverted megaphone-shaped vocal tract with less mouth opening and more laryngeal space has been related to classical singing and head voice, as well as to 'soft voice' supported by closed vowels, such as [i:] and [u:].[3] However, for a deeper understanding on the connections between different voice productions guided by different terminologies and pedagogical concepts, further studies are needed.

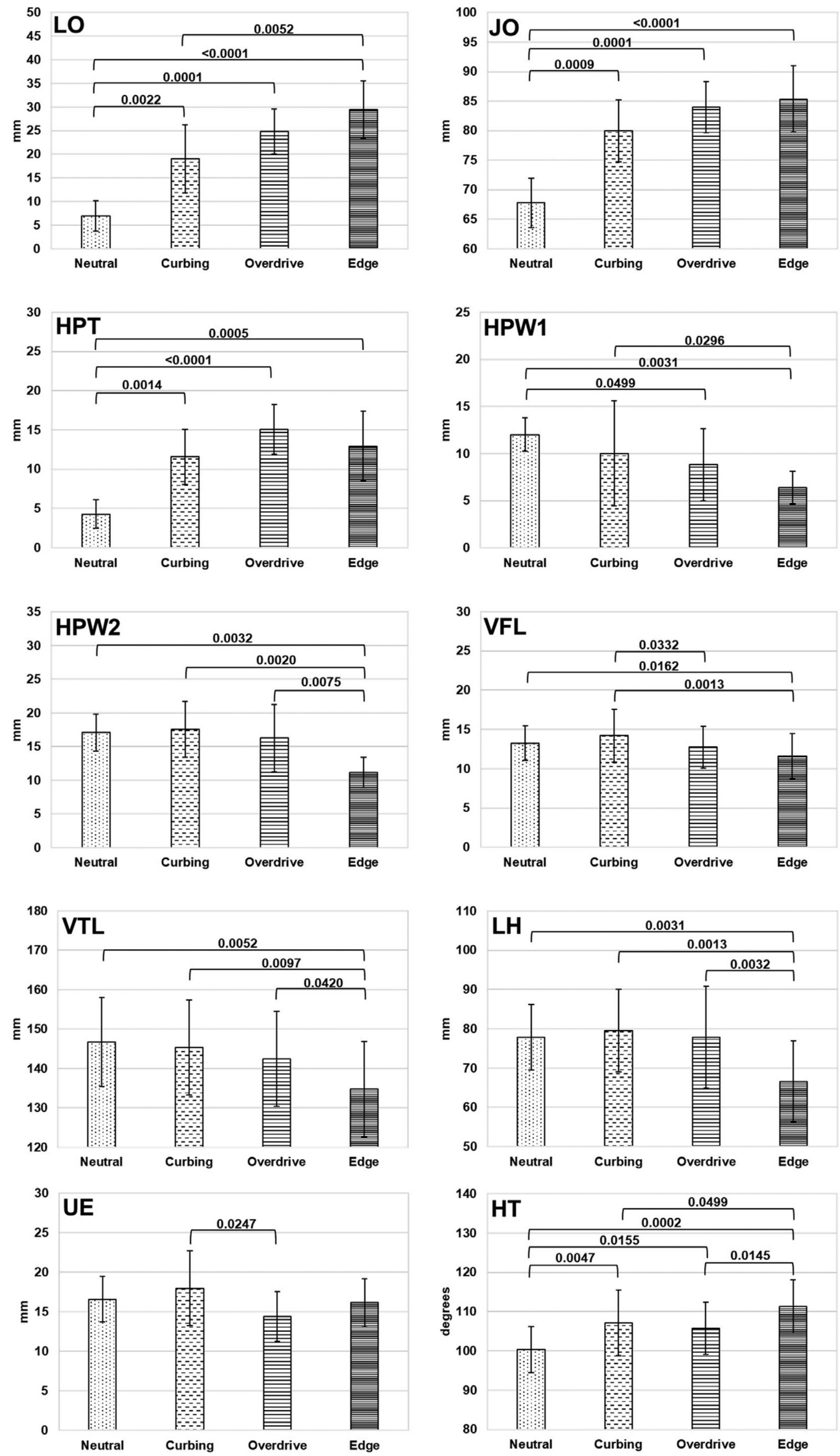

FIGURE 4. Intermodal differences between the measured vocal tract dimensions of the vocal modes presented as absolute mean measurement values with standard deviation and *P* values. Statistical analysis was performed to support the observed trends on the data.

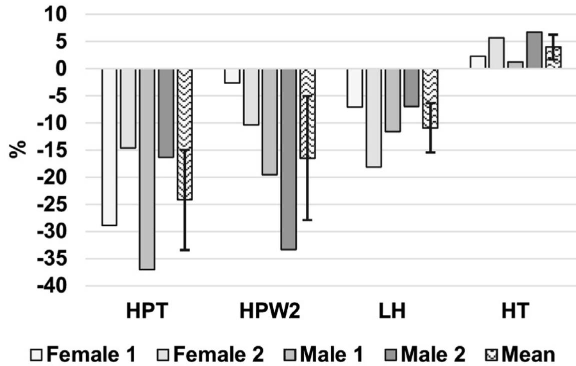

FIGURE 5. Changes from Overdrive to Edge using same vowel [ɛ:]. Differences between the vocal tract configurations were found in the shortest distance between tongue and palate (HPT), the shortest distance between epiglottis and pharyngeal wall just above the laryngeal inlet (HPW2), LH and head tilt (HT). Absolute measurement values were compared as changes in percentages. Similar changes from Overdrive to Edge were found for all singers.

It is noteworthy, that nonmetallic Neutral differs from the metallic vocal modes in not being dependent on vowels.[38] Thus, vocal tract shapes used in the metallic vocal modes might be possible in Neutral. In our data, closed vowels [i:] and [u:] had been chosen for Neutral, since they are recommended as practicing vowels for nonmetallic sounds[38] and in general, singers find it easy to produce Neutral with these vowels. However, since Curbing, Overdrive and Edge are restricted to certain vowels and vowel variations, it could be concluded that certain amount of mouth opening, laryngeal rise and narrowing, relative to Neutral, is needed when singing in the metallic vocal modes, and certain vowels provide vocal tract positions that enable these adjustments. Much work, however, is needed in order to shed light on how vowels and different voice characteristics interact, and how these auditory dimensions are related to vocal tract adjustments.

Variations in VFL were observed in all singers. Curbing was found to differ from Edge by showing longer vocal folds (VFL) and less anterior-posterior narrowing (HPW2). Compared to Edge, also Neutral showed some tendencies for relatively longer vocal folds accompanied by less anterior-posterior narrowing (HPW2). Similar findings on VFL have been provided by several studies on different voice categorizations. According to X-ray observations, vocal folds have been reported to be shorter in 'sharp voice' and longer in 'soft voice', while 'ordinary voice' was found to fall in between these two.[3] In videofiberscopic observations, shorter vocal folds have been found for louder sounds compared to quieter sounds.[62] In a radiographic study, longest vocal folds were detected in 'piano covered' singing compared to 'forte open' and 'forte covered' singing.[18] Based on endoscopic observations, laryngeal anterior-posterior constriction accompanied by shorter vocal folds have been described for belting when compared to 'head' and 'mix', 'mix' being in between and 'head' displaying the longest vocal folds.[28] In a laryngostroboscopic study, anterior laryngeal tilting with lengthening of the vocal folds were visually observed for reduced density when compared to fuller density in relation to Curbing, Overdrive and Edge.[43] Endoscopy, however, does not allow full visibility to the vocal folds due to anterior-posterior narrowing, and thus cannot provide accurate data on differences in VFL. Variation in VFL on same pitch phonation in the four vocal modes is presumably related to changes in vocal fold oscillatory characteristics, contributing to the perception of the vocal modes. Further studies on vocal fold dynamics in the four vocal modes are needed to confirm such connections.

Head tilt was not part of characterizing vocal tract dimensions. Yet, in our study, it seemed to play a role in the vocal mode production. Edge stood out by showing the most and Neutral the least posteriorly tilted head position compared to other vocal modes. According to previous MRI studies, some classical male singers have shown to curve the cervical spine at higher pitches resulting in anterior movement of the head accompanied by increasing lip and jaw opening.[19,48] Head tilt could be related to different vocal functions,[54] yet the possible connection to vocal tract adjustments regarding vocal modes remain unclear.

Intramodal differences

Lower vocal tract adjustments displayed differences between the vocal modes but also within the vocal modes due to two

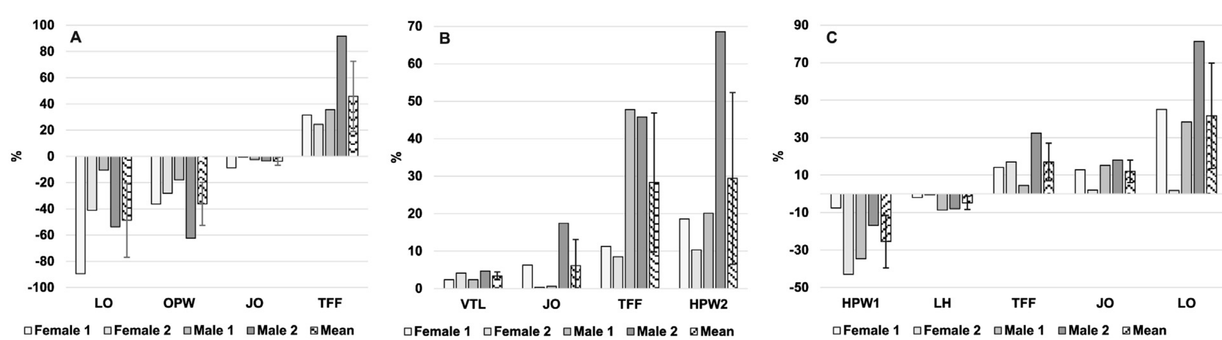

FIGURE 6. Intramodal differences. Differences in the measured vocal tract dimensions between two vowels representing the same vocal mode as changes in percentages for each singer with mean and standard deviation. A. Changes from Neutral [i:] to [u:]. B. Changes from Overdrive [ɛ:] to [ɔ:]. C. Changes from Edge [ɛ:] to [æ:].

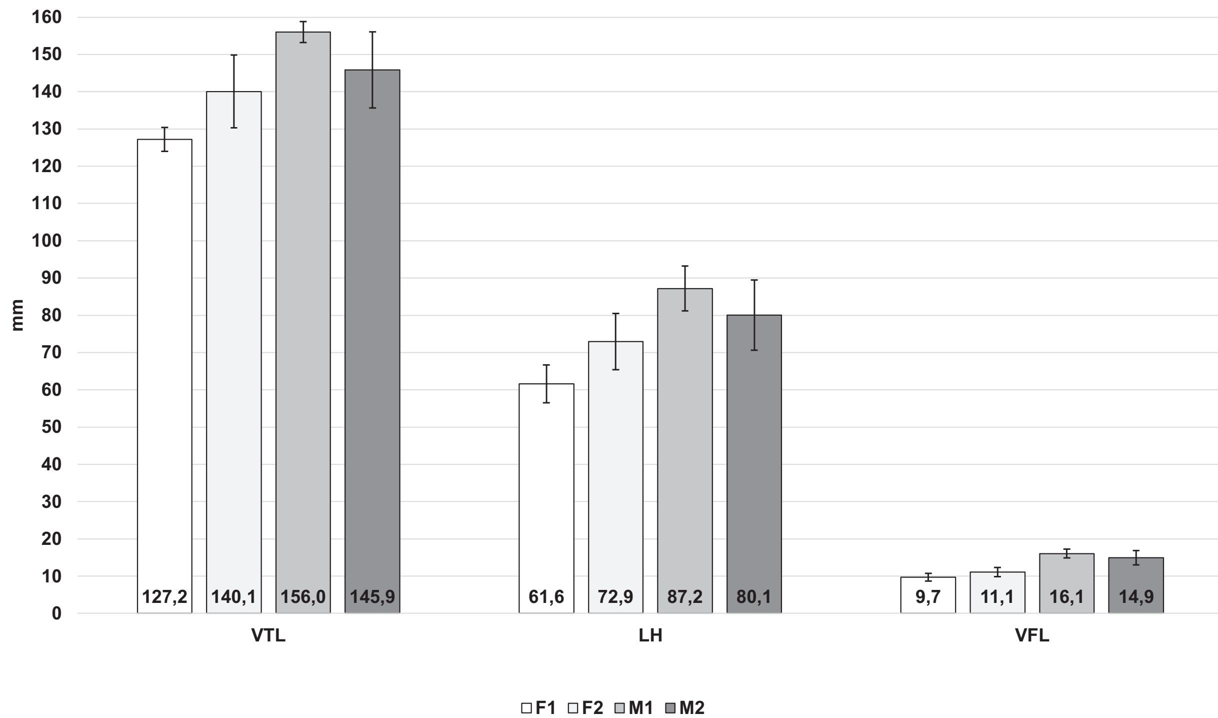

FIGURE 7. Singer-specific review. Average values of VTL, LH and VFL displayed differences between the singers.

different vowels used in each mode. Similarly, in previous studies, lower vocal tract positioning has been found to possess great importance for voice quality, and further, lower vocal tract measurements have been found to vary with vowel quality.[25,63] Moreover, our findings on tongue position are in line with the 'vowel shapes' described in literature.[1-4] Yet, interindividual differences were also detected. Thus, more studies on individual strategies in vowel production[64] related to vocal modes, are called for.

Interindividual differences

Vocal tract size-related measurements were calculated from the data to see the possible differences between the singers. The results showed consistent differences regarding VTL, LH and VFL measurements, according to which a clear order was found. F1 showed lowest measured values for all three measurements followed by F2, M2 and finally M1. These interindividual differences correspond to general differences between genders[2] as well as findings on vocal tract

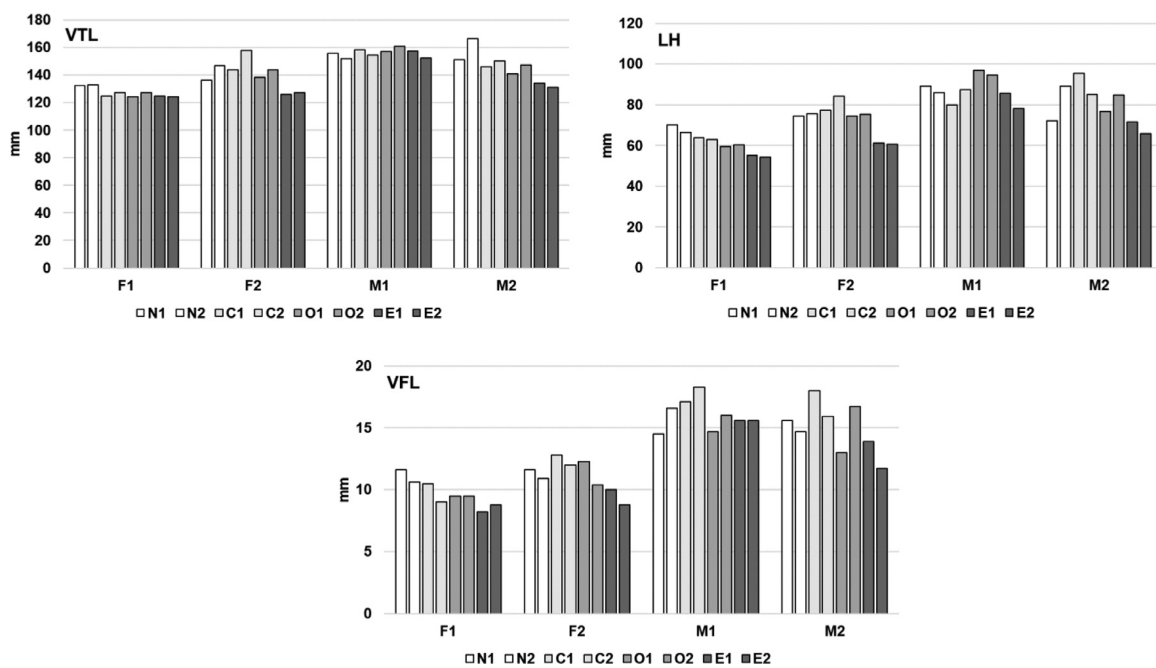

FIGURE 8. Absolute measured values of VTL, LH and VFL for each singer. For labels N1, N2, C1, C2, O1, O2, E1 and E2 (Table 1).

sizes in relation to voice classification.[29,65] However, the measurements were done from images taken during sound production. To include measurements from resting position of the vocal tract, axial plane should be included in the image processing; sagittal projection cannot provide accurate data on vocal fold level due to abduction of the vocal folds.

Regarding LH and VFL, Edge did not stand out from other vocal modes for M1 as clearly as for other singers. Moreover, F1 was shown to decrease VFL, raise the larynx (LH) and shorten the vocal tract (VTL) for Curbing, moving closer to Overdrive and Edge settings, while for other singers such consistent behaviour was not found. These findings may be related to observations made during auditory evaluation, according to which differences between singers were detected regarding vocal mode characteristics. Edge samples from M1 were perceived as 'softer' and 'less metallic' compared to other singers, while F1 was observed to sound 'sharper' and 'more metallic' especially in Curbing compared to other singers. Although no generalization can be drawn from these observations, they provide an interesting perspective for future research regarding individual morphology and articulatory manoeuvres in relation to differences in auditory perception.[29]

Limitations

The current study has several limitations. The number of participants was low. This is, however, a common issue with all MRI work in the field of voice research. The vocal mode samples were given in only one pitch and the vowel selection was limited. In this study, however, the purpose was to investigate vocal tract configurations in such vowel conditions that are recommended when practicing each vocal mode. Further, the aim was to examine what these 'practicing vowels' offer for each vocal mode, how Neutral differs from the metallic vocal modes in regards to the selected vowels, and what is the difference between Curbing, Overdrive and Edge in connection to the vowels that work in these vocal modes. Examining Neutral with the vowels used in Curbing, Overdrive and Edge, would offer a more comprehensive basis in the comparison between the vocal modes, and thus, more MRI studies are needed to form a deeper understanding of the vocal tract behaviour in the production of the vocal modes.

Limitations regarding MRI methodology are acknowledged. While imaged in the MRI scanner, the singers are exposed to loud MRI noises and gravitational effects due to supine position. Isolating headphones interfere with auditory perception and voice production and the audio signal is collared by the method used to capture the vocal output. However, the aim of this study was to investigate vocal tract behaviour in the four vocal modes proposed, described and used in the CVT method. Since vocal tract behaviour is controlled by the singer, trained singer's own somatosensory feedback of the vocal mode performance played a key role in the real time evaluation procedure.[55,56] For untrained subjects, differences in the measured vocal tract dimensions has been detected in supine position compared to upright position, due to gravitational effect.[54] In turn, only minor differences were found for professional tenors, respectively.[66] In our study, such comparison between supine and upright position was not possible to perform since scanner enabling upright imaging setup was not available. However, the singers were able to practice and perform the vocal tasks in a supine position in the scanner and all samples were recorded under the same gravitational conditions.

The segmentation work was performed by the first author with the knowledge of the vocal modes being used in each sagittal projection. In this respect, data analysis was performed in unblinded manner. However, the anatomical accuracy of manual segmentations was supervised by a highly experienced neuroradiologist without any prior experience of different singing techniques or the scientific field of voice.

Mid-sagittal projection aims to illustrate the midsection of the vocal tract, reflecting most of the relevant anatomical structures and anterior-posterior dimensions of the vocal tract cavities, while ignoring other important factors. The soft tissue around the arytenoid cartilages appear differently in the sagittal projection depending on the selection from the 3D image stack. For a more detailed examination, coronal and axial planes should be included. Such review, however, was out of the scope of this pilot study.

Finally, this study could not provide reliable measured values for the shortest distances (ELT, LT) due to insufficient image resolution of 1.4 mm.

No generalizations can be drawn from this study. Also, for investigating the interdependence between measured parameters, larger material is required.

CONCLUSIONS

Vocal tract adjustments play a key role in the production of the vocal modes. The model used to measure vocal tract dimensions succeeded in finding significant differences between the vocal modes, also detecting differences between different vowel productions. When examining different voice characteristics and categorizations, individual voices and morphologies, and the way they work in these concepts, should be considered. The method used to characterize vocal tract dimensions seem promising and would be worthwhile to apply to a larger material.

Acknowledgements

We warmly thank our volunteer singers and assistants for their patience and availability. We also thank Ahmed Geneid, Anne-Maria Laukkanen, Jenni Holappa, Michael Keitel, Marianne Haapea, Tero Ikävalko, Cathrine Sadolin, Mathias Aaen, Julian McGlashan and several CVT teachers for valuable advices and fruitful discussions on various aspects on human voice and research methodology. Furthermore, the corresponding author gratefully acknowledges The Education Fund, Finnish Cultural Foundation North Ostrobothnia Regional Fund, The

University of Oulu Scholarship Foundation, and Oulu University of Applied Sciences, The Department of Media and Performing Arts, for supporting this work.

SUPPLEMENTARY DATA

Supplementary data related to this article can be found online at https://doi.org/10.1016/j.jvoice.2020.01.015.

REFERENCES

1. Sundberg J. *The Science of the Singing Voice*. DeKalb, IL: Northern Illinois University Press; 1987.
2. Titze IR. *Principles of Voice Production*. 2nd ed Iowa City: National Center for Voice and Speech; 2000.
3. Chiba T, Kajiyama M. *The Vowel, Its Nature and Structure*. Tokyo: Phonetic Society of Japan; 1958.
4. Fant G. *Acoustic Theory of Speech Production: With Calculations Based on X-Ray Studies of Russian Articulators*. 'S-Gravenhage: Mouton; 1960.
5. Rossing TD. *The Science of Sound*. 2nd ed Reading, MA: Addison-Wesley; 1990.
6. Titze IR, Story BH. Acoustic interactions of the voice source with the lower vocal tract. *J Acoust Soc Am*. 1997;101:2234–2243. https://doi.org/10.1121/1.418246.
7. Titze IR. Nonlinear source−filter coupling in phonation: theory. *J Acoust Soc Am*. 2008;123:2733–2749. https://doi.org/10.1121/1.2832337.
8. Titze IR, Worley AS. Modeling source-filter interaction in belting and high-pitched operatic male singing. *J Acoust Soc Am*. 2009;126: 1530–1540. https://doi.org/10.1121/1.3160296.
9. Titze IR, Abbott KV. *Vocology: The Science and Practice of Voice Habilitation*. Salt Lake City, UT: National Center for Voice and Speech; 2012.
10. Garellek M, Samlan R, Gerratt BR, et al. Modeling the voice source in terms of spectral slopes. *J Acoust Soc Am*. 2016;139:1404–1410. https://doi.org/10.1121/1.4944474.
11. Titze IR, Bergan CC, Hunter EJ, et al. Source and filter adjustments affecting the perception of the vocal qualities twang and yawn. *Logop Phoniatr Vocol*. 2003;28:147–155. https://doi.org/10.1080/14015430310018874.
12. Appelman DR. *The Science of Vocal Pedagogy: Theory and Application*. Bloomington: Indiana University Press; 1986.
13. *Handbook of the International Phonetic Association: A Guide to the Use of the International Phonetic Alphabet*. Cambridge: Cambridge University Press; 1999. The reference is published and authored by "International Phonetic Association". Please see: https://www.bookdepository.com/Handbook-of-the-International-Phonetic-Association-International-Phonetic-Association/9780521637510?redirected=true&utm_medium=Google&utm_campaign=Base2&utm_source=FI&utm_content=Handbook-of-the-International-Phonetic-Association&selectCurrency=EUR&w=AFFFAU9SG75S95A8V7NV&pdg=pla-317692435101:cmp-8750579716:adg-88443198215:crv-410017274306:pos-:dev-c&gclid=Cj0KCQiAkKnyBRDwARIsALtxe7igi6t_rTsSPGY03znbJSnVPxYzEIJq8OhfXUmn5RdXSyb1E2AQGHgaAgYSEALw_wcB.
14. Peterson GE, Barney HL. Control methods used in a study of the vowels. *J Acoust Soc Am*. 1952;24:175–184. https://doi.org/10.1121/1.1906875.
15. Sonninen A, Hurme P, Vilkman E. Roentgenological observations on vocal fold length-changes with special reference to register transition and open/covered voice. *Scand J Logop Phoniatr*. 1992;17:95–106. https://doi.org/10.3109/14015439209098719.
16. Hollien H. *On Vocal Registers*. Gainesville: Communication Sciences Laboratory, Florida University; 1972.
17. Titze IR, Alipour F. *The Myoelastic Aerodynamic Theory of Phonation*. Denver, CO: National Center for Voice and Speech; 2006.
18. Sonninen A, Hurme P, Laukkanen A-M. The external frame function in the control of pitch, register, and singing mode: radiographic observations of a female singer. *J Voice*. 1999;13:319–340. https://doi.org/10.1016/S0892-1997(99)80039-3.
19. Echternach M, Sundberg J, Arndt S, et al. Vocal tract and register changes analysed by real-time MRI in male professional singers—a pilot study. *Logop Phoniatr Vocol*. 2008;33:67–73. https://doi.org/10.1080/14015430701875653.
20. Echternach M, Popeil L, Traser L, et al. Vocal tract shapes in different singing functions used in musical theater singing—a pilot study. *J Voice*. 2014;28:653.e1–653.e7. https://doi.org/10.1016/j.jvoice.2014.01.011.
21. Henrich Bernardoni N, Smith J, Wolfe J. Vocal tract resonances in singing: variation with laryngeal mechanism for male operatic singers in chest and falsetto registers. *J Acoust Soc Am*. 2014;135:491–501. https://doi.org/10.1121/1.4836255.
22. Hanayama EM, Camargo ZA, Tsuji DH, et al. Metallic voice: physiological and acoustic features. *J Voice*. 2009;23:62–70. https://doi.org/10.1016/j.jvoice.2006.12.006.
23. Duvvuru S, Erickson M. The effect of change in spectral slope and formant frequencies on the perception of loudness. *J Voice*. 2013;27:691–697. https://doi.org/10.1016/j.jvoice.2013.05.004.
24. Guzman M, Lanas A, Olavarria C, et al. Laryngoscopic and spectral analysis of laryngeal and pharyngeal configuration in non-classical singing styles. *J Voice*. 2015;29:130.e21–130.e28. https://doi.org/10.1016/j.jvoice.2014.05.004.
25. Mainka A, Poznyakovskiy A, Platzek I, et al. Lower vocal tract morphologic adjustments are relevant for voice timbre in singing. Bolhuis JJ, ed. *PLoS One*. 2015;10: e0132241. https://doi.org/10.1371/journal.pone.0132241.
26. Echternach M, Burk F, Burdumy M, et al. Morphometric differences of vocal tract articulators in different loudness conditions in singing. Bolhuis JJ, ed. *PLoS One*. 2016;11: e0153792. https://doi.org/10.1371/journal.pone.0153792.
27. Sundberg J, Gramming P, Lovetri J. Comparisons of pharynx, source, formant, and pressure characteristics in operatic and musical theatre singing. *J Voice*. 1993;7:301–310. https://doi.org/10.1016/S0892-1997(05)80118-3.
28. Lovetri J, Lesh S, Woo P. Preliminary study on the ability of trained singers to control the intrinsic and extrinsic laryngeal musculature. *J Voice*. 1999;13:219–226. https://doi.org/10.1016/S0892-1997(99)80024-1.
29. Roers F, Mürbe D, Sundberg J. Voice classification and vocal tract of singers: a study of x-ray images and morphology. *J Acoust Soc Am*. 2009;125:503–512. https://doi.org/10.1121/1.3026326.
30. McGlashan J, Thuesen MA, Sadolin C. Overdrive and Edge as refiners of "Belting"? *J Voice*. 2017;31:385.e11–385.e22. https://doi.org/10.1016/j.jvoice.2016.09.006.
31. Bourne T, Garnier M. Physiological and acoustic characteristics of the female music theater voice. *J Acoust Soc Am*. 2012;131:1586–1594. https://doi.org/10.1121/1.3675010.
32. Bourne T, Garnier M, Samson A. Physiological and acoustic characteristics of the male music theatre voice. *J Acoust Soc Am*. 2016;140:610–621. https://doi.org/10.1121/1.4954751.
33. Sundberg J, Thalén M, Popeil L. Substyles of belting: phonatory and resonatory characteristics. *J Voice*. 2012;26:44–50. https://doi.org/10.1016/j.jvoice.2010.10.007.
34. Fadel CBX, Dassie-Leite AP, Santos RS, et al. Acoustic characteristics of the metallic voice quality. *CoDAS*. 2015;27:97–100. https://doi.org/10.1590/2317-1782/20152014159.
35. Kob M, Henrich N, Herzel H, et al. Analysing and understanding the singing voice: recent progress and open questions. *Curr Bioinform*. 2011;6:362–374. https://doi.org/10.2174/157489311796904709.
36. López-Íñiguez G, Pozo JI. Analysis of constructive practice in instrumental music education: case study with an expert cello teacher. *Teach Teach Educ*. 2016;60:97–107. https://doi.org/10.1016/j.tate.2016.08.002.
37. Sansom R. Voice teacher certification and research: relationships and trends. *Voice Speech Rev*. 2019;13:127–129. https://doi.org/10.1080/23268263.2019.1625589.

38. Sadolin C. *Complete Vocal Technique*. Denmark: Shout Publishing; 2017. iTunes App Store © & Google Play Store © Copenhagen.
39. Brixen EB, Sadolin C, Kjelin H. On acoustic detection of vocal modes. *132nd Convention of the Audio Engineering Society*. Budabest, Hungary; 2012.
40. Brixen EB, Sadolin C, Kjelin H. Acoustical characteristics of vocal modes in singing. *134th Convention of the Audio Engineering Society*. Rome, Italy; 2013.
41. Thuesen MA, McGlashan J, Sadolin C. Curbing—the metallic mode in-between. *J Voice*. 2017;31:644.e1–644.e10. https://doi.org/10.1016/j.jvoice.2017.01.010.
42. Brixen EB, Sadolin C, Kjelin H. The importance of onset features in listeners' perception of vocal modes in singing. *137th Convention of the Audio Engineering Society*. Los Angeles,CA; 2014.
43. Mathias A, McGlashan J, Sadolin C. Investigating laryngeal "tilt" on same-pitch phonation—preliminary findings of vocal mode metal and density parameters as alternatives to cricothyroid-thyroarytenoid "mix.". *J Voice*. 2018. https://doi.org/10.1016/j.jvoice.2018.02.023.
44. Sundberg J, Bitelli M, Holmberg A, et al. The "Overdrive" mode in the "Complete Vocal Technique": a preliminary study. *J Voice*. 2017;31:528–535. https://doi.org/10.1016/j.jvoice.2017.02.009.
45. Echternach M, Markl M, Richter B. Vocal tract configurations in yodelling—prospective comparison of two Swiss yodeller and two non-yodeller subjects. *Logop Phoniatr Vocol*. 2011;36:109–113. https://doi.org/10.3109/14015439.2011.566576.
46. Echternach M, Sundberg J, Baumann T, et al. Vocal tract area functions and formant frequencies in opera tenors' modal and falsetto registers. *J Acoust Soc Am*. 2011;129:3955–3963. https://doi.org/10.1121/1.3589249.
47. Echternach M, Sundberg J, Arndt S, et al. Vocal tract in female registers—a dynamic real-time MRI study. *J Voice*. 2010;24:133–139. https://doi.org/10.1016/j.jvoice.2008.06.004.
48. Echternach M, Sundberg J, Markl M, et al. Professional opera tenors' vocal tract configurations in registers. *Folia Phoniatr Logop*. 2010;62:278–287. https://doi.org/10.1159/000312668.
49. Tom K, Titze IR, Hoffman EA, et al. Three-dimensional vocal tract imaging and formant structure: varying vocal register, pitch, and loudness. *J Acoust Soc Am*. 2001;109:742–747. https://doi.org/10.1121/1.1332380.
50. Story BH, Titze IR, Hoffman EA. The relationship of vocal tract shape to three voice qualities. *J Acoust Soc Am*. 2001;109:1651–1667. https://doi.org/10.1121/1.1352085.
51. Baer T, Gore JC, Gracco LC, et al. Analysis of vocal tract shape and dimensions using magnetic resonance imaging: vowels. *J Acoust Soc Am*. 1991;90:799–828. https://doi.org/10.1121/1.401949.
52. Echternach M, Birkholz P, Sundberg J, et al. Resonatory properties in professional tenors singing above the passaggio. *Acta Acust United Acust*. 2016;102:298–306. https://doi.org/10.3813/AAA.918945.
53. Guzman M, Laukkanen A-M, Krupa P, et al. Vocal tract and glottal function during and after vocal exercising with resonance tube and straw. *J Voice*. 2013;27:523.e19–523.e34. https://doi.org/10.1016/j.jvoice.2013.02.007.
54. Traser L, Burdumy M, Richter B, et al. Weight-bearing MR imaging as an option in the study of gravitational effects on the vocal tract of untrained subjects in singing phonation. Larson CR, ed *PLoS One*. 2014;9: e112405. https://doi.org/10.1371/journal.pone.0112405.
55. Mürbe D, Pabst F, Hofmann G, et al. Effects of a professional solo singer education on auditory and kinesthetic feedback—a longitudinal study of singers' pitch control. *J Voice*. 2004;18:236–241. https://doi.org/10.1016/j.jvoice.2003.05.001.
56. Zarate JM. The neural control of singing. *Front Hum Neurosci*. 2013;7. https://doi.org/10.3389/fnhum.2013.00237.
57. Pätynen J, Katz BFG, Lokki T. Investigations on the balloon as an impulse source. *J Acoust Soc Am*. 2011;129:EL27–EL33. https://doi.org/10.1121/1.3518780.
58. Hertegard S, Hakansson A, Thorstensen O. Vocal fold length measurements with computed tomography. *Scand J Logop Phoniatr*. 1993;18:57–63. https://doi.org/10.3109/14015439309101350.
59. Hannila I, Lammentausta E, Tervonen O, et al. The repeatability of T2 relaxation time measurement of human knee articular cartilage. *Magn Reson Mater Physics Biol Med*. 2015;28:547–553. https://doi.org/10.1007/s10334-015-0494-3.
60. Lecler A, Savatovsky J, Balvay D, et al. Repeatability of apparent diffusion coefficient and intravoxel incoherent motion parameters at 3.0 Tesla in orbital lesions. *Eur Radiol*. 2017;27:5094–5103. https://doi.org/10.1007/s00330-017-4933-6.
61. Estill J. Belting and classic voice quality: some physiological differences. *Med Probl Perform Art*. 1988;3:37–43.
62. Lindestad P-Å, Södersten M. Laryngeal and pharyngeal behavior in countertenor and baritone singing—a videofiberscopic study. *J Voice*. 1988;2:132–139. https://doi.org/10.1016/S0892-1997(88)80069-9.
63. Echternach M, Traser L, Richter B. Vocal tract configurations in tenors' passaggio in different vowel conditions—a real-time magnetic resonance imaging study. *J Voice*. 2014;28:262.e1–262.e8. https://doi.org/10.1016/j.jvoice.2013.10.009.
64. Johnson K, Ladefoged P, Lindau M. Individual differences in vowel production. *J Acoust Soc Am*. 1993;94:701–714. https://doi.org/10.1121/1.406887.
65. Roers F, Mürbe D, Sundberg J. Predicted singers' vocal fold lengths and voice classification—a study of x-ray morphological measures. *J Voice*. 2009;23:408–413. https://doi.org/10.1016/j.jvoice.2007.12.003.
66. Traser L, Burdumy M, Richter B, et al. The effect of supine and upright position on vocal tract configurations during singing—a comparative study in professional tenors. *J Voice*. 2013;27:141–148. https://doi.org/10.1016/j.jvoice.2012.11.002.