Factors of the assessment of a vocal category in female singers - a preliminary study

M. Frič and A. Kulanová
Musical acoustics research cent, Academy of performing arts in P, Malostranske nam.13, 118 00 Praha, Czech Republic
marekfric@centrum.cz

Recordings of 11 women of different vocal category were evaluated using listening tests on examples of recitatives and coloraturas from Rossini's aria "Una voce poco fa". Seven voice teachers (trained group) and seven musicians non-teachers (untrained group) evaluated the properties of the vocal category, timbre (dark-light), the resonance, the vowel position, suitability of vibrato, aesthetic impression, and vocal mobility. The results showed a significantly higher interjudge reliability (interclass correlation) in the trained group. The highest reliability was achieved at timbre and vocal category evaluation, the least consistent was the evaluation of the resonance. Factor analysis of the assessment variability showed dependent rating of the vocal category, lightness and vowel position for both groups in recitative. The trained group similarly evaluated the lightness and the vocal category in coloratura. Assessment of the vocal category correlated with the reported categories of singers only in the trained group. LTA spectra differentiate the groups of mezzo-sopranos and sopranos in the spectral band of 3.5-4 kHz in recitative and in the band of 1.5-3.5 and 4.5-6.5 kHz in coloratura. The singers, who were able to sing for a longer time in a higher position in coloratura and have lower values of the first formant amplification, were evaluated as a higher vocal category and lighter voice. The lightness was associated with a higher position of the center of gravity (COG) in the band of 2-5 kHz, the vocal category with an overall spectral width given by a higher position of COG in 5-10 kHz.

1 Introduction

The traditional division of the vocal categories of classical opera singers comes from the voice properties such as pitch position, pitch range and the timbre. The classification into specific categories in the praxis carried out on the basis of experience of the singer's voice teacher.

From an objective point of view knowledge of the acoustic properties of vocal categories are based mainly on the results of studies dealing with the perceptual evaluation of sung vocals [1-3]. These results showed that perceptual differences in timbre decreases with an increasing pitch, the timbral differences are greater between categories than between individuals of the same category, and the ability to distinguish between different singers from each other based on vowels comparison decreases with the pitch difference of tones [4].

Acoustic studies have shown differences between vocal categories at the pitch of the voice [4], the characteristics of the voice range profile [5-9], and in the formant positions [6], in particular in the spectral band of the singer's formant [1,10,11]. The maximum intensity of this band is different among categories [1] and also differentiates the singing performance levels in professional sopranos [12].

These observations confirm the physiological conditions of the differences among the categories based on the length of the vocal folds and vocal tract proportions [13,14]. However, these observations have not been linked directly to listening evaluation of the real vocal performances.

The aim of this study is to verify the ability to distinguish and identify the vocal categories on real recordings of arias, to compare the perceptual ratings of singing voice between trained (experienced) and untrained group, and to interpret these assessments in relation to acoustic measurements.

2 Materials and methods

2.1 Stimuli

Freely available recordings of the Rossini's aria "Una voce poco fa" from the opera Il Barbiere di Seviglia in the key of E major were selected for listening tests. As the reference of vocal category of stimuli were taken biographical data of singers (7 mezzo-sopranos: 2 dramatic, 2 lyric, 2 coloratura and 1 unspecified, 4 sopranos: 1 dramatic, 1 lyric, 1dramatic coloratura and 1 coloratura).

The recordings were cut into two examples as listening stimuli. The first type of stimuli (**recitative**) contained the initial three and a half bars (approximately 10 seconds) of recitative "Una voce poco fa quinel cor miri suono." The second type of stimuli (**coloratura**) contained the text "a Lindoro mio sara" in the tone range of E4-G#5.

2.2 Listening tests

Two types of stimuli (recitative and coloratura) were evaluated for the 6 properties of the singing voice (see Table 1). The degree of resonance and the evaluation of light - dark timbre were performed using ranking and scaling listening test without repeating of the evaluation. Other properties were evaluated using categorical scaling test in two consecutive repetitions (test-retest).

Fourteen listeners participated in the listening tests, which were divided according to their profession in untrained and trained group. Trained group consisted of seven trained solo singers, they were also vocal pedagogues. Seven listeners participated in untrained group, all of them were musicians but anybody was neither singing teacher nor professional singer.

2.3 Acoustic analysis

The stimuli recordings were analyzed in 100 ms length window with the step of 10ms. For each segment were determined F0, SPL, LPC and Fourier spectra, centre of spectral gravity (COG), position and amplification of formants, singing power ratio. For whole stimuli recordings were estimated long time average spectra and LPC spectra.

2.4 Data analysis

The final rating for each stimulus in each group was calculated as the average value in the case that the rating of a particular stimulus for the property had a normal distribution (Lilliefors test). Otherwise, the median was taken as the resulting value.

To evaluate the inter-rater agreement was used value ICC1 coefficient (Cronbach's alpha).

Factor analysis was used to determine the interrelationships of the final evaluation of properties in groups. It searches for the variability factors in the evaluation of 6 properties separately for each task and each group. Spearman's correlation, ICC1 and ICC2 were used to determine the correlation between the resulting assessments of vocal category and reported category.

Table 1: Rated properties of singing voice for the types of stimuli (R – recitative, C – coloratura) and number of trials. Listening methods RS – ranking and scaling, CS – categorical scaling (labels of categories are listed in columns 1-7).

Stimulus	Property	Listening method	Num. of trials	1	2	3	4	5	6	7
R/C	Timbre	RS	1	light						dark
R/C	Resonance	RS	1	little						much
R/C	Vocal category	CS	2	alt	dram. mezzo	lyr. mezzo	col. mezzo	dram. sop	lyr. sop	col. sop
R/C	Vibrato	CS	2	slow	mild	adequate	fast	tremolo	extreme	
R/C	Vowel position	CS	2	extr. front	front	adequate	back	extr. back		
R	Aesthetic impression	CS	2	bad	inadequate	neutral	adequate	good	beautiful	excellent
C	Voice mobility	CS	2	immobile	hardly	moderate	good	naturally coloratura		

According to biographical data the singers were divided into the groups of mezzo-sopranos and sopranos. The average values of the measured acoustic parameters were compared by mean of t-test. Average values of acoustic parameters were correlated using Spearman's correlation with the results of categorical scaling tests and using Pearson's correlation with the results of ranking and scaling tests (resonance and light-dark timbre).

3 Results

3.1 Perceptual evaluation

The interclass correlation results revealed a moderate to high consensus for all evaluators.

In recitative, the assessment of the vibrato was the most consistent (0.953), followed by evaluation of dark-light timbre (0.928). In the evaluation of coloratura, the most consistent was the assessment of timbre (0.923), followed by the vocal category (0.919). The evaluation of resonance was the less consistent in both recitative (0.732) and coloratura (0.531).

The comparison of consistency between trained an untrained group revealed higher ICC1 values in all evaluated properties except vowel position in coloratura for trained group. This difference in consistency between groups was significant (paired t-test, $p < 0{,}005$).

3.2 The evaluation of vocal category

The resulting evaluation of the vocal category for both recitative and coloratura significantly correlated (Table 2) with bibliographical data of subject only in trained group. All correlation coefficients (Spearman, ICC1 and ICC2) were much higher for the evaluation of the trained group.

3.3 Factor analysis of resulting evaluation

The resulting factors (Table 3) combine properties which assessment related to each other. In the evaluation of recitative in both untrained and trained listeners were similarly assessed properties lightness, the voice category and the position of vocals. In the evaluation of recitative in the second factor untrained group similarly evaluated the resonance and the vibrato; the aesthetic impression was an independent third factor. In the evaluation of trained group were separately evaluated the aesthetic impression as the second factor, and vibrato separately as the third factor.

In coloratura evaluation, the untrained group evaluated lightness, vocal category and position of vowels in the first factor. The second factor contained vibrato, the third factor consisted only from the voice mobility. The trained group evaluated together in the first factor lightness and vocal category, in the second factor were similarly evaluated vibrato and the voice mobility. The resonance created the third factor of the evaluation.

Table 2: Correlation of the vocal category rating results and reported data. Spearman – rank correlation, PVAL – level of statistical significance, ICC1 – Cronbach's alpha, ICC2 – intra class correlation Type 2.

Stim.	Recitative		Coloratura	
Group	Untrained	Trained	Untrained	Trained
Spearman	0.251	0.755	0.354	0.77
PVAL	0.456	0.007	0.286	0.006
ICC1	0.174	0.843	0.33	0.841
ICC2	0.095	0.728	0.198	0.726

Table 3: The results of the factor analysis of the resulting assessment of the vocal properties for recitative and coloratura, respectively. Bold letters indicate significant correlation between the factor and evaluated property.

Group	**Untrained**			**Trained**		
Recitative	**Fa1**	**Fa2**	**Fa3**	**Fa1**	**Fa2**	**Fa3**
Timbre (light-dark)	**-0.93**	-0.23	0.08	**-0.86**	-0.33	-0.38
Resonance	0.39	**0.87**	0.00	0.36	***0.67***	0.54
Vocal category	**0.82**	0.42	-0.01	**0.96**	0.07	0.12
Vibrato	0.15	**0.95**	0.13	0.23	0.08	**0.95**
Vowel position	**-0.83**	-0.12	0.43	**-0.74**	-0.51	-0.30
Aesthetic impression	0.16	-0.11	**-0.97**	0.15	**0.94**	0.02
Cumul. var.%	**56.5**	**82.1**	**91.8**	**65.6**	**80.4**	**92.8**
Coloratura	**Fa1**	**Fa2**	**Fa3**	**Fa1**	**Fa2**	**Fa3**
Timbre (light-dark)	**0.78**	0.23	0.16	**0.87**	0.03	-0.33
Resonance	-0.41	***-0.69***	-0.02	-0.13	0.24	**0.94**
Vocal category	**-0.95**	-0.01	-0.01	**-0.96**	0.01	0.05
Vibrato	-0.30	**0.87**	0.07	0.17	**0.93**	0.11
Vowel position	**0.89**	-0.33	-0.19	***0.62***	-0.24	***-0.68***
Voice mobility	0.02	0.05	**0.99**	-0.34	**0.75**	0.39
Cumul. var.%	**42.7**	**67.8**	**83.5**	**53.8**	**79.6**	**89.3**

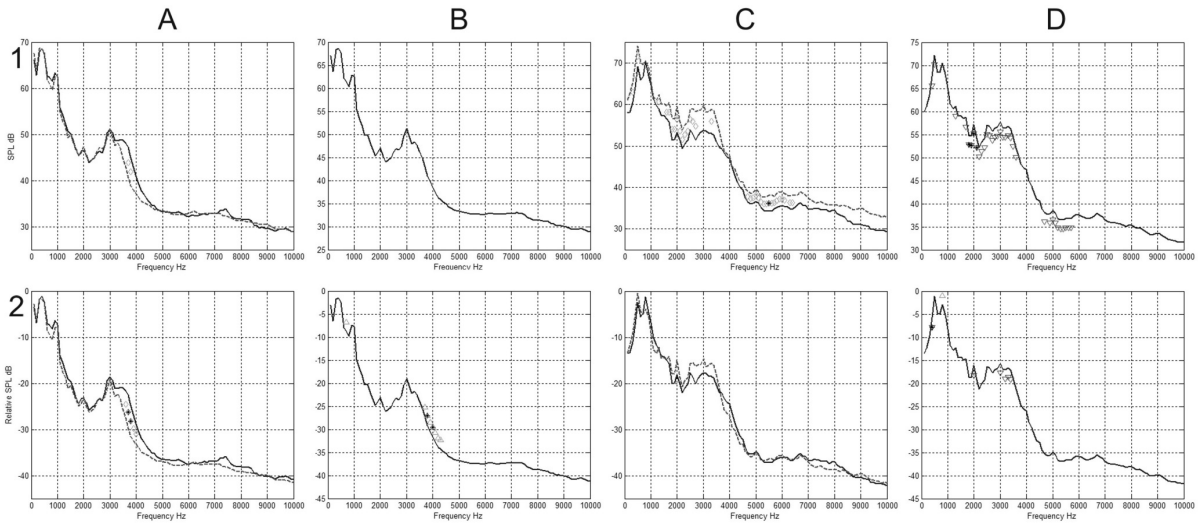

Figure 1: Comparison of long-term averaged spectra (LTAS) (1st row absolute value, the second row relative values to the highest average level) between the sopranos (solid dark line) and mezzo-soprano (dashed light line) in recitative (column A) and coloratura (column C). Correlation between reported vocal categories and LTA spectral bands (columns B and D), upwards triangles indicate positive trend correlation, downwards triangles - negative correlation for recitative (column B) and coloratura (column D), respectively. ◊ - indicates the trend difference between vocal categories, * - indicated bands wit statistically significant levels after Bonferroni correction.

3.4 Comparison of acoustic properties

Comparison of relative LTA spectra between the soprano and mezzo-soprano groups (

Figure 1) shows the trend differences in the spectral band of 3.5-4 kHz in recitative, where sopranos achieved higher absolute and relative levels. This band positively correlated with evaluation of vocal category with higher tessitura. In the coloratura, there were trend differences in the spectral band of 1.5-2.5 kHz and 4.5-6.5 kHz, where mezzos reached higher absolute levels. The correlation between reported vocal category and absolute LTAS shows negative trend correlation ($p<0.05$) in the similar spectral bands as the difference mentioned above. In coloratura, the relative LTAS shows negative correlation with category in the band 0.5 kHz and 3-3.5 kHz and the positive correlation in the band of second formant (0.8 kHz).

Comparison of the average values of acoustic parameters for all measured segments between reported groups of sopranos and mezzo-sopranos (Table 4) showed higher values of the first formant amplification (A1) and the centre of spectral gravity (COG) in the band 2-5 kHz in mezzo in recitative. In coloratura there were statistical difference in values of singing power ratio (SPR) and the position of the fourth formant (F4). SPR was higher in mezzo and F4 was higher in soprano voices.

3.5 Correlation among results of evaluated properties and acoustical parameters

Table 5 shows the most significant correlations among measured parameters and properties of the vocal category (referred and evaluated) and timbre (light – dark).

The amplitude of the first formant (A1) was the best correlated parameter with reported vocal category, as in the recitative so in coloratura. In perceptual evaluation, however, A1 correlated only in coloratura with both voice category and the lightness. The centre of gravity in the spectral band 2-5 kHz (COG 2-5 kHz) and the position of the spectral maximum in the band 2-4 kHz (FSH) similarly correlated with the evaluation of lightness in trained group both in recitative and coloratura. Untrained group connected this parameter with the evaluation of the vocal category.

Table 4: Comparison (t-test) of acoustic parameters means between soprano and mezzo-soprano group. SPR – Singing power ratio, F4 position of the fourth formant, A1 – first formant amplification, COG 2-5 kHz – centre of spectral gravity in the band of 2-5 kHz.

	Parameter	SPR	F4	A1	COG2-5kHz
Recitative	sopranos	-24.97	3342.5	25.80	3085.0
	std	2.10	111.5	2.89	90.4
	mezzos	-23.93	3222.9	29.75	2947.1
	std	2.07	98.4	2.21	71.8
	pval	0.443	0.097	**0.031**	**0.021**
Coloratura	sopranos	-19.51	3332.5	26.56	2967.5
	std	2.55	108.7	1.29	62.9
	mezzos	-15.03	3171.4	26.13	2888.6
	std	2.98	76.0	1.60	143.5
	pval	**0.033**	**0.017**	0.662	0.331

For coloratura, the average voice pitch was important in the evaluation of vocal category in the untrained group, whereas in the trained group the pitch was associated with the lightness. The position of the first formant (F1) correlated with the reported categories only in the untrained group. Sound pressure level (SPL) was negatively related to the reported vocal category. The value SPR correlated negatively with the lightness and vocal category.

Table 5: Correlation among average values of measured acoustical parameters and reported vocal category and results of listening tests of the timbre (light – dark) and vocal category for recitative and coloratura, respectively and for trained and untrained group of listeners. All displayed values indicated statistical significance p<0.05, * - p<0.01, ** - p<0.005.

Stimul	Property / Parameter	Group	Pitch	SPL	SPR	F1	A1	COG25k	FSH	LPC Pmax 0-1.61kHz	LPC COG 1.6-5kHz	LPC COG5-10kHz	Spectral end
Recitative	Reported voc. cat.	X					-0.86**						0.81*
	Vocal category	Untrained						0.69	0.61				
	Light - dark												
	Vocal category	Trained											0.67
	Light - dark							-0.64	-0.65				
Coloratura	Reported voc. cat.	X		-0.64	-0.71		-0.76*						0.75
	Vocal category	Untrained	0.87**			0.71	-0.74*		0.65	0.65			
	Light - dark				0.69		0.66				-0.61	-0.79*	
	Vocal category	Trained			-0.78*		-0.65					0.64	0.88**
	Light - dark		-0.81*		0.67		0.67	-0.64	-0.62	-0.74x	-0.64		

From the LPC analysis of the whole stimuli, the position of the maximum of the spectral band of first two formants (LPC Pmax 0-1.6 kHz) correlated with the evaluation of the vocal category in coloratura in untrained group, whereas in trained group with lightness. The COG of LPC spectra in the band 1.6-5 kHz (LPC COG 1.5-5 kHz) correlated with lightness in coloratura in both groups. COG in the band 5-10 kHz correlated with lightness in untrained group and with the vocal category in trained group. Subjective evaluated of the end of LTA spectra correlated with the reported data and evaluation of the vocal category in trained group in both recitative and coloratura.

4 Discussion

The present study confirms the assumption of the generally more consistent evaluation of voice properties in the trained group than in untrained. Only the assessment of the vocal category by trained group achieved significant degree of agreement with the biographical reported data of singers. The evaluation of timbre (lightness) belonged to the most consistent in both groups of listeners. In coloratura the trained group the most consistently evaluated the lightness and untrained group the vocal category, respectively. These results confirm that the property of the timbre light-dark belongs to the basic perceptual dimension of the perceptual evaluation of sounds so the consistency of this property doesn't depend on the specific listening training for the voice. The assessment of resonance was generally the worst consistent evaluation. A possible conflict in the evaluation of resonance can be attributed to different interpretations and also multidimensionality of this property [15].

Factor analysis confirmed the coherence of the assessment of the vocal category and lightness. In recitative there was also associated the evaluation of the vowel position with these properties, whereas in coloratura in trained group the assessment of the vowel position was better related to the degree of resonance. Untrained group linked the assessment of resonance and vibrato; however, trained group evaluated the adequacy of vibrato together with the mobility of the voice.

The comparison of average values of acoustical parameters for all segments of stimuli revealed higher values of the first formant amplification in mezzo-soprano voices in recitative and higher values of singing power ratio (the level difference of spectral maxima between 2-4 kHz and 0-2 kHz) in coloratura. According to [16], where the level of the bands 0.8 and 2.5 kHz were perceptually associated with the full, melodious and colourful voice, we can think of the darker and full timbre of mezzo-sopranos.

On the other hand, sopranos had higher position of the COG 2-5 kHz in recitative and higher position of F4 in coloratura. These data were shown similarly in the comparison of LTA spectra of recitative, where the trend of higher levels of the band 3.5-4 kHz in sopranos was confirmed. That correspond to the assumption of the higher position of the higher (singer's) formant in sopranos [1,10,11,17]. Negative correlation of the levels in the band 2.5-3.5 kHz in coloratura with the vocal category verifies the presupposition of the lower levels of singer's formant, respective its absence in higher soprano voices [11]. Similarly to [2] the position of the centre of gravity in the band 2-5 kHz differed mezzo from soprano voices.

The first formant amplification correlated with the reported vocal category of subject in both recitative and coloratura but in perceptual evaluation it was true only in

coloratura. This parameter was connected similarly with dark timbre in coloratura evaluation in both groups.

The average pitch was significant only in coloratura, where untrained group connected the ability to sing at higher positions for a longer time with the higher vocal category, whereas the trained group it connected with the lighter voice.

From the whole LPC spectra, the position of COG in the band 1.6-5 kHz was generally connected with lightness. While the COG 5-10 kHz was in coloratura associated with lightness in untrained group and with vocal category in trained group, the opposite was the true for the position of maximum in 0-1.6 kHz LPC spectral band. These results indicate that some acoustic properties are oppositely connected with properties of lightness or vocal category between trained and untrained groups.

5 Conclusion

The study confirmed the better consistency in assessment of voice properties in the trained group of singing teachers and also their ability to better determine the voice category of subjects. In general, the evaluation of timbre light-dark was highly consistent, at least the assessment of the degree of resonance. Factor analysis showed that the evaluation of vocal categories related to the lightness and in the recitative also to vowel position. Based on reported vocal categories of subject, sopranos and mezzo-sopranos differed in position of maximum and the level of the spectral band 2-4 kHz, in recitative also in the amplification of first formant. In perceptual assessment were found the relationships between lightness and the position of maximum and spectral centroid of 2-4 kHz band. Only in coloratura the amplification of the first formant was related to both vocal category and darkness, the position of maximum LPC spectra in the band 0-1.6 kHz was related to lightness and the COG 5-10 kHz to vocal category. The presented relationships, in some cases, show the confusion assessment lightness and voice category between trained and untrained group.

Acknowledgments

Supported by the Ministry of Education, Youth and Sports of the Czech Republic in the Long Term Conceptual Development of Research Institutes grant of the Academy of Performing Arts in Prague: The "Sound quality" project.

References

[1] G. Bloothooft and R. Plomp, "Spectral analysis of sung vowels. III. Characteristics of singers and modes of singing," *J.Acoust.Soc.Am.* 79 (3): 852-864 (1986).

[2] M.L. Erickson, "Dissimilarity and the classification of female singing voices: a preliminary study," *J Voice* 17 (2): 195-206 (2003).

[3] M.L. Erickson, "Dissimilarity and the classification of male singing voices," *J Voice* 22 (3): 290-299 (2008).

[4] S.I. Chernobel'skii, "[The speaking fundamental frequency and the singing voice type]," *Vestn.Otorinolaringol.* (6): 36-37 (2010).

[5] T.F. Cleveland, "Acoustic properties of voice timbre types and their influence on voice classification," *J Acoust.Soc.Am.* 61 (6): 1622-1629 (1977).

[6] A.M. Johnson and G.B. Kempster, "Classification of the classical male singing voice using long-term average spectrum," *J Voice* 25 (5): 538-543 (2011).

[7] H. Lycke, W. Decoster, A. Ivanova, M.M. Van Hulle, and F.I. de Jong, "Discrimination of three basic female voice types in female singing students by voice range profile-derived parameters," *Folia Phoniatr.Logop.* 64 (2): 80-86 (2012).

[8] H. Lycke, A. Ivanova, M.M. Van Hulle, W. Decoster, and F.I. de Jong, "Discrimination of three basic male voice types by voice range profile-derived parameters," *Folia Phoniatr.Logop.* 65 (1): 20-24 (2013).

[9] Ch.T. Herbst, E. Duus, H. Jers, and J.G. Švec, "Quantitative Voice Class Assessment of Amateur Choir Singers: A Pilot Investigation," *International Journal of Research in Choral Singing (IJRCS)* 4 (1): 47-59 (2012).

[10] G. Berndtsson and J. Sundberg, "Perceptual significance of the center frequency of singer's formant," *STL-QPSR* 35 (4): 95-105 (1994).

[11] J. Sundberg, "Level and center frequency of the singer's formant," *J Voice* 15 (2): 176-186 (2001).

[12] J.J. Barnes, P.F. Davis, J.F. Oates, and J. Chapman, "The relationship between professional operatic soprano voice and high range spectral energy," *J.Acoust.Soc.Am.* 116 (1): 530-538 (2004).

[13] F. Roers, D. Murbe, and J. Sundberg, "Predicted singers' vocal fold lengths and voice classification-a study of x-ray morphological measures," *J Voice* 23 (4): 408-413 (2009).

[14] F. Roers, D. Murbe, and J. Sundberg, "Voice classification and vocal tract of singers: a study of x-ray images and morphology," *J Acoust.Soc.Am.* 125 (1): 503-512 (2009).

[15] J.M. Oates, B. Bain, P. Davis, J. Chapman, and D. Kenny, "Development of an auditory-perceptual rating instrument for the operatic singing voice," *J.Voice* 20 (1): 71-81 (2006).

[16] G. Bloothooft and R. Plomp, "The timbre of sung vowels," *J Acoust.Soc.Am.* 84 (3): 847-860 (1988).

[17] M.L. Erickson, "The interaction of formant frequency and pitch in the perception of voice category and jaw opening in female singers," *J Voice* 18 (1): 24-37 (2004).