HOW DOES STUDYING INFLUENCE ONE'S VOICE QUALITY?

Allan Vurma and Tarmo Pajusaar
Estonian Academy of Music, Tallinn, Estonia
Einar Meister
Institute of Cybernetics, Tallinn, Estonia
Jaan Ross
Institute of Estonian Language, Tallinn, Estonia

ABSTRACT

Excerpts from voices of 42 students at the Estonian Academy of Music, with different length of their study, were digitally recorded and acoustically analyzed with respect to fundamental and formant frequencies. It was found that objective characteristics of a student's voice often does not meet accepted standards of a good-quality opera voice. The level of the singer's formant and the amplitude of frequency vibrato tend to increase systematically with the number of years studied.

INTRODUCTION

Formants are peaks of spectral envelope caused by the vocal tract resonances. While the first two formants are mostly responsible for differentiation of different vowels, the function of higher formants is believed to influence the voice color. The so-called singer's formant is a spectral peak at the frequency region between 2500 and 3000 Hz, observed predominantly in spectra of male voices. It enables an opera singer to be heard "over" the big symphony orchestra because the spectral envelope of orchestral sound decreases more rapidly than that of a trained singer and also because of higher sensitivity of the human auditory system in the region specified above.

The singer's formant is considered to be an acquired skill. Its articulatory origin is believed to be in lowering the larynx as well as widening the pharynx and the Morgan cavities inside the larynx box which results in clustering together the 3rd and the 4th formants [1]. The singer's formant is an important factor in the perception of a singer's voice category. Higher voices tend to have the singer's formant at higher frequencies [2].

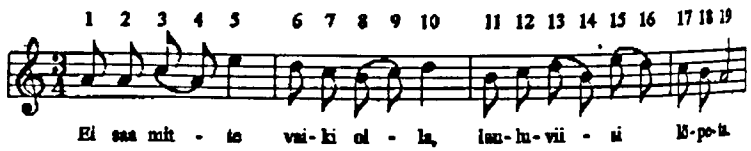

Figure 1. The phrase from a song by Estonian composer Miina Härma, "Ei saa mitte vaiki olla".

MATERIAL AND MEASUREMENTS

A seven-word phrase from the beginning of a well-known melody by Estonian composer Miina Härma (see Fig. 1) was recorded three times: as sung in E minor or A minor, or as spoken, by 42 voice students (16 male and 26 female) from the Estonian Academy of Music. The duration of years studied before the recording was different for individual students, and varied from 1 to 9 years. The students have been instructed by different professors (total of 12). Recordings were made with a DAT recorder (SONY TCD-D3) and a microphone ML-19 in a room with small reverberation.

All recordings were subsequently rated by 4 experts (3 professional singers and 1 musicologist) on two scales. Tone quality and intonation purity were both rated on a five-point scale. In addition, experts were asked to determine the category of voice (tenor, baritone, or bass for males, soprano or mezzo-soprano for females) a singer should belong to, according to the recording.

Measurements were made with a Kay Elemetrics Company Computerized Speech Laboratory (CSL), with a sampling rate of 16 kHz. For each singer, long-term average spectra (Hamming window, 1024 points) were computed for the whole phrases, and LPC (filter order of 16, frame length 15 ms, autocorrelation method) formant histories as well as fundamental frequency fluctuations for the separate syllables were found.

RESULTS

For the voices studied, it was possible to distinguish four different shapes of the spectral envelope in the region of the singer's formant, i.e. around 2.7 kHz (see Fig. 2): (1) a clearly pronounced triangle, (2) rise of the envelope which may result either in a rather flat maximum or in two equally strong peaks, (3) a small rise of the envelope with 3 to 5 equally strong peaks at the maximum (occurs mostly in female singers), and (4) no clear level increase. Table 1 presents the following average data separately for 2 sexes, 2 tonalities, and 5 voice categories: singer's formant frequency, its level, and the normalized frequency distance between the 3rd and the 4th formants.

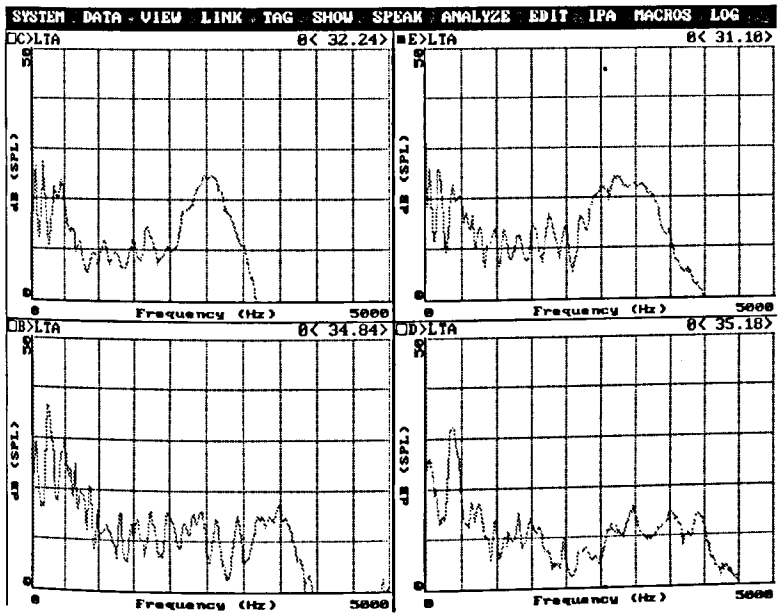

Figure 2. Four different shapes of the singer's formant: a triangle (top left), a flat maximum (top right), with no clear level increase (bottom left), and with a couple of smaller peaks (bottom right).

Analysis of variance demonstrates influence on the level of the singer's formant by (1) the normalized frequency distance between the 3rd and the 4th

formants ($p<0.001$), (2) the duration of a singer's vocal studies ($p<0.01$), and (3) the singer's sex ($p<0.001$). Greater distance between the 3rd and the 4th formants corresponds to smaller level of the singer's formant. The level increases with the number of years studied: e.g., the average for female singers of 1 to 2 years of study is -20.5 dB with respect to the highest peak in the spectrum, in comparison to -15.6 dB for those of 7 and more years of study. The level is generally 10 dB higher in male than in female voices. Intrasex differences in the level are not significant.

The frequency of the singer's formant is significantly different in male and female voices ($p<0.05$). The singer's formant tends to be higher in frequency, higher is the voice, providing that a mezzosoprano follows a tenor in ascending rank order. Intrasex differences in the frequency, however, are not significant. The expert ratings of the voice quality significantly correlates neither with the level nor the frequency of the singer's formant.

Table 1. Averaged data on the singer's formant frequency (fm2, Hz), level (sform, dB), and the normalized distance between the 3rd and the 4th formants [(F4-F3)/F3], with standard deviations (σ).

	key	Σ	fm2	σ	sform	σ	(F4-F3)/F3	σ
fem	E	26	3337	569	-18.4	3.5	38	6.2
fem	A	26	3065	560	-18.4	3.7	43	6.3
male	E	16	2717	338	-8.4	4.9	26	5.5
male	A	13	2732	304	-6.1	8.9	25	5.5
sopr	E	18	3456	572	-17.7	3.4	39	6.7
sopr	A	18	3165	592	-17.9	3.7	44	6.2
mezzo	E	8	3127	555	-20.0	3.4	36	4.8
mezzo	A	8	2840	429	-19.5	3.7	39	4.5
tenor	E	6	2612	339	-10.1	7.5	27	6.3
tenor	A	6	2914	359	-6.2	12.4	26	6.6
barit	E	6	2734	359	-7.2	2.4	24	6.3
barit	A	5	2585	154	-5.0	5.6	24	5.4
basso	E	4	2848	344	-7.8	2.4	23	2.2
basso	A	3	2552	33	-7.5	6.3	23	0.3

LPC formant history analysis was used predominantly in order to compare differences between voice production results for speech and singing. However, because of high fundamental frequency and resulting big frequency distances between harmonics, the LPC analysis results for lower formants tend to coincide with the harmonic frequencies. The results therefore may be used for higher formant frequencies only.

In singing, all formant frequencies are less in value than in speech. The most pronounced differences between different voice categories seem to be in F3 for male speech ($p<0.05$) and in F5 for female singing voices ($p<0.05$).

For the frequency vibrato, its rate and amplitude have so far been measured for two notes (syllables), nos. 5 and 19 (see Fig. 1). The vibrato rate varies from 4.8 to 9.9 Hz for the investigated singers. This variation across the singers is considered too wide, as the normally accepted limits of the vibrato rate are between 5.5 and 7.5 Hz [3]. Its large variation for our subjects can be understood as evidence about the untrained nature of their voices. The amplitude of vibrato was between 9 to 113 cents which, according to [3], is within acceptable limits of up to 200 cents. There is a significant tendency for the vibrato rate to increase with the higher voice category ($p<0.01$) as well as a significant tendency for the vibrato amplitude to increase with the duration of study ($p<0.001$). It is interesting to notice that in comparing vibrato of voice students at 4 different professors, those studying at LT had the vibrato amplitude almost twice as much as the rest of the students ($p<0.01$).

As to experts' recording-based decisions about the students' voice category, the analysis of variance shows significant dependence between the two ($p<0.001$). Confusions may occur when one and the same singer performs in the lower (E minor) and the higher tonality (A minor), and in the former case is classified as a mezzosoprano but in the latter case as a soprano.

DISCUSSION AND CONCLUSIONS

The present data show a tendency for the singer's formant to increase in level as one's vocal studies proceed. At the same time, however, we did not find correlation between experts' ratings of the tone quality and the level of the singer's formant. In many cases, the quality of a student's voice did not meet the standards of a good opera voice. The diffuse peak in the spectrum corresponding to the singer's formant may indicate certain inconsistency in voice production technique.

It must be pointed out that male and female singers must undertake different amount of effort to concentrate energy at the frequency of the singer's formant. Male voices tend to possess weaker fundamental in their glottal signal than female [3]. While in female singing voices the fundamental tends to be matched to the first formant, in male voices there is a match rather between the 2nd or the 3rd harmonic and the first formant. Consequently, there may be problems of matching the singer's formant to a certain harmonic for female voices, since the frequency distance between the harmonics is more sparse there. The strong fundamental may be characteristic to female voices in general: it has been claimed [4] that tenors deliberately bypass the production of a strong fundamental in order to avoid the female quality of their voice.

REFERENCES

[1] Sundberg, J. (1974), "Articulatory interpretation of the "singing formant"," *J. Acoust. Soc. Am.*, vol. 55, pp. 838-844.

[2] Dmitriev, L. and Kiselev, A. (1979), "Relationship between the formant structure of different types of singing voices and the dimension of supraglottal cavities," *Folia Phoniatrica*, vol. 31, pp. 238-241.

[3] Sundberg, J. (1987), *The Science of the Singing Voice*, Illinois: DeKalb.

[4] Titze, I. R., Mapes, S., and Story, B. (1994), "Acoustics of the tenor high voice," *J. Acoust. Soc. Am.*, vol. 95, pp. 1133-1142.