POWER CONTROL FOR THE SECOND HARMONIC AT HIGH PITCHES IN SOPRANO SINGING:
A CASE STUDY

Hironori Takemoto
National Institute of Information and Communications Technology
takemoto@nict.go.jp

Seiji Adachi
Fraunhofer Institute Seiji.Adachi@ibp.fraunhofer.de

Takeshi Saitou
Kanazawa University t-saitou@ec.t.kanazawa-u.ac.jp

Kiyoshi Honda
Tianjin University khonda@sannet.ne.jp

Eri Haneishi
Showa University of Music haneishi@tosei-showa-music.ac.jp

Hiroko Kishimoto
Showa University of Music kisimoto@tosei-showa-music.ac.jp

Sopranos have many singing skills relating to control of the vocal tract shape, especially at high pitches. In the present study, one of the skills called girare in Italian opera was examined. This skill is said to be used for producing a mild voice at high pitches. The midsagittal images of the vocal tract of a soprano were scanned by magnetic resonance imaging during singing of the vowel /a/ at G5 (784 Hz) and A5 (880 Hz) with and without girare. Acoustic analysis showed that girare selectively depressed the power of the second harmonic by more than 10 dB. This result implied that the power reduction could make the auditory impression of singing voice mild. Image analysis indicated that girare elongated the vocal tract, widened the oral cavity, and slightly constricted the lips. According to acoustic sensitivity analysis, all these changes tend to decrease the frequency of the second resonance of the vocal tract. Acoustic simulation confirmed that these changes in the vocal tract successfully reduced the power of the second harmonic by more than 10 dB.

1. INTRODUCTION

It is well known that sopranos tune the vocal tract resonances (R1, R2, etc.) to the voice harmonics (f0, 2f0, etc.) especially at high pitches. The way of tuning was investigated in detail by direct measurements of the vocal tract resonances and the voice harmonics using broadband excitation at the mouth [1, 2]. Garnier et al. [1] revealed how sopranos change the relationships among R1, R2, f0, and 2f0 depending on the pitch. On the other hand, apart from the lips it is still unclear how sopranos control the vocal tract shape to effect changes in the vocal tract resonances, because the vocal tract is an invisible internal space. To overcome this problem, magnetic resonance imaging (MRI) could be a useful tool. If the skill in controlling the vocal tract shape could be visualized, the physiological and acoustic features of the skill would be revealed.

Recently, we measured the vocal tract shape for sopranos using MRI during singing to examine how sopranos increase R1 at high pitches [3]. In that study, the vocal tract area functions at three different pitches were extracted from MRI data. From these area functions, transfer functions and functions of acoustic sensitivity [4, 5] to R1 were calculated and examined. As a result, it was revealed that sopranos tended to change the area and length of selected parts in the vocal tract which had high sensitivity to R1. In short, sopranos effectively changed the vocal tract shape to increase R1 at high pitches.

In the present study, we focused on one of the skills in soprano singing, called girare in Italian opera. According to informal interviews with several Japanese sopranos, this skill contributes to producing a mild voice at high pitches, and it could be similar to the skill called “cover” in English opera and “decken” in German opera. Although these sopranos were very interested in how they controlled the vocal tract when singing with girare, it is difficult for most of them to sing both with and without girare, because they have trained to sing with girare for a long time. One of them, however, offered to participate in the acoustic and MRI experiments as a subject, because this soprano could sing with and without girare at limited pitches for several seconds. Thus, in this preliminary study, first we analyzed the subject’s singing voice and measured midsagittal images of the vocal tract by MRI during singing with and without girare. Then, we examined how the soprano changed the vocal tract when using girare. Furthermore, we discussed the acoustic function of this skill based on acoustic sensitivity analysis.

2. MATERIALS AND METHODS

2.1 Subjects and Task

The subject was a lyric soprano who was in her 50s and professional for 27 years. The subject was asked to sustain the pitches of G5 (784 Hz) and A5 (880 Hz) on the vowel /a/ with and without girare. To allow natural singing, no restriction was placed on the use of vibrato.
2.2 MRI Experiment and Sound Recording

The midsagittal images of the vocal tract of the subject during singing at the two pitches with and without girare were scanned by MRI (Siemens MAGNETOM Verio 3T) installed in the Brain Activity Imaging Center at ATR Promotions Inc. The subject took a supine position in the MRI gantry and was asked to keep singing approximately for 6 s. During singing, MRI scanning for 5 s was performed. The pixel resolution was 1.0 mm, the slice thickness was 7.0 mm, the echo time was 2.43 ms, and the repetition time was 20 ms. Because the MRI scanner generated loud noise (over 90 dB), the singing voices for the acoustic analysis were recorded by a microphone (Denmark Pro Audio, DPA 4066) and a solid state recorder (Marantz, PMD 671) in an anechoic room just after the MRI experiment. In the recording, the subject took a supine position in the same way as in the MRI experiment.

2.3 Computational Simulation

The area functions of the vocal tract at G5 and A5 for the subject could not be obtained because image distortion occurred around a dental filling in lateral planes. Instead, the area function at G5 for another soprano [3] was used to estimate acoustic effects of changes in the vocal tract caused by girare, because the overall shape of the vocal tract for the vowel /a/ might be regarded as common among sopranos. Based on the estimated acoustic effects, physiological and acoustic features of girare were simulated by a method proposed by Adachi et al. [6] which can modify the vocal tract based on acoustic sensitivity functions.

3. RESULTS AND DISCUSSION

3.1 Acoustic Effects of girare

Figures 1 and 2 show long-term spectra at G5 (784 Hz) and A5 (880 Hz) with (red) and without (blue) girare. Both figures indicate that the spectrum with girare was smaller in power than that without girare in the frequency range approximately from 1 to 3 kHz. Especially, the power of 2f0 reduced remarkably when using girare: a reduction of approximately 10 dB at G5, and 13 dB at A5. It was possible that this remarkable power reduction of 2f0 made the auditory impression mild. Note that all three sopranos could discriminate singing voices with and without girare at both pitches in the informal listening test.

There are possibly two factors for such a remarkable power reduction: changes in the vocal tract resonance and those in the vocal fold vibration pattern. Because non-harmonic components and the overall spectral tilt which reflected the vocal fold vibration pattern were common between the two spectra with and without girare at the same pitch, the vocal fold vibration pattern probably did not change much. Thus, changes in the vocal tract resonance around 2f0 could have caused the power reduction.

As R2 existed in the frequency region from 1 to 2 kHz which included 2f0 at G5 and A5, it was reasonable to consider that R2 changes affected the power of 2f0. According to Garnier et al. [1], around these two pitches, G5 and A5, sopranos tend to tune R1 to f0 and R2 to 2f0, while individual differences also exist. The tuning prominently enhances the power of both f0 and 2f0. If R2 were not tuned to 2f0, the power of 2f0 would reduce remarkably. Thus, it is hypothesized that girare is a skill to intentionally avoid tuning R2 to 2f0 at pitches where R1 and R2 are normally tuned to f0 and 2f0, respectively. At this moment, however, it is unclear whether the soprano increased or decreased R2 for detuning, while keeping R1 at f0.

![Figure 1](image1.png)

**Figure 1.** Power spectra for the vowel /a/ sung with (red) and without (blue) girare at G5 (784 Hz).

![Figure 2](image2.png)

**Figure 2.** Power spectra for the vowel /a/ sung with (red) and without (blue) girare at A5 (880 Hz).

3.2 Changes in the vocal tract shape caused by girare

Figures 3 and 4 show outlines of the speech organs extracted from the midsagittal images of the vowel /a/ sung with (red) and without (blue) girare at G5 (784 Hz) and A5 (880 Hz). At both pitches, the speech organs were displaced by girare in the same way: the tongue, the lower jaw, the hyoid bone, and the larynx were drawn in the postero-inferior direction and the upper and lower lips were slightly protruded and constricted. These displacements would elongate the vocal tract, widen the oral cavity, and constrict the lips. In the following, we will discuss acoustic effects of these changes in the vocal tract shape on the vocal tract resonance.

Figure 5b presents an artificial area function which emulates the vocal tract for the vowel /a/ sung at G5 without girare. This area function was obtained by modifying an area function for another soprano [3] to tune R1 to f0 (G5: 784 Hz) and R2 to 2f0 (double frequency of G5: 1568 Hz) using a resonance tuning method proposed by Adachi et al. [6]. In this method, small changes in area and length were iteratively applied by referring to area and length sensitivity functions. Note that technically R1 of this area function was tuned to slightly above G5, 800 Hz, because of stable oscillation of the vocal folds.
3.3 Simulation of girare

First, we verified whether the vocal tract modification for decreasing R2 could cause similar changes in the vocal tract shape to those estimated from MRI data. In this simulation, the area function shown in Fig. 5b was used as the original vocal tract shape, i.e., the one without girare. As described in the previous section, R1 and R2 of this vocal tract were tuned to 800 Hz and 1568 Hz, respectively. To emulate a vocal tract shape with girare, the resonance tuning method of Adachi et al. [6] was applied again to the original area function in Fig. 5b, modifying it to maintain R1 at 800 Hz while decreasing R2 from 1568 Hz to 1200 Hz.

Figure 6a presents the original (blue) and modified (red) area functions. As the figure shows, the modification widened the oral cavity, and both constricted and elongated the lip region. These changes agreed with those estimated from the midsagittal images of the vocal tract shown in Figs. 3 and 4. Thus, the vocal tract modification for decreasing R2 successfully simulated the physiological features of girare.

Figure 6b shows transfer functions calculated from the original (blue) and modified (red) area functions discussed above. As R2 decreased, the power at 1568 Hz decreased by approximately 20 dB. Thus, if f0 is set to G5 (784 Hz) and girare is used, the power of the second harmonic at 2f0 (1568 Hz) would reduce by approximately 20 dB, theoretically.

To examine the power reduction, we simulated vowel production using these area functions and a two-dimensional model of vocal fold vibration [7]. The reason why a two-dimensional model was used for this simulation rather than the two-mass model [8] was that the two-dimensional model can oscillate at high pitches close to R1 where the two-mass model is difficult to oscillate. In this simulation, when the vocal fold resonance frequency was set to 720 Hz and the subglottal pressure was set to 1500 Pa, f0 became G5 (784 Hz). Other parameters were set to default values of the two-dimensional model.

Figure 7 shows long-term spectra at G5 with (red) and without (blue) girare: i.e., the red spectrum was calculated from the modified area function, while the blue spectrum was calculated from the original area function.
Although the spectral tilt is different from that of the soprano voice because the two-dimensional model cannot simulate falsetto register, the power of 2f0 for the red spectrum was approximately 12 dB lower than that for the blue spectrum. In short, the vocal tract modification for decreasing R2 suppressed the power of 2f0 by 12 dB. This value was close to the observed one described in Sec.3.1. Thus, the vocal tract modification for decreasing R2 quantitatively simulated the acoustic feature of girare.

Figure 6. (a) original (blue) and modified (red) area functions, and (b) corresponding transfer functions.

Figure 7. Vowel power spectra calculated by long-term Fourier analysis of vowel waveforms synthesized from the original (blue) and modified (red) area functions shown in Fig. 6 (a).

4. CONCLUSION

In the present study, we examined the mechanism of a soprano skill, called girare. The acoustic recording and MRI experiments for a soprano were performed during singing with and without girare at G5 (784 Hz) and A5 (880 Hz). At both pitches, the power of 2f0 reduced by more than 10 dB when the subject used girare. According to midsagittal images of the vocal tract, it was estimated that the vocal tract was elongated, the oral cavity was widened, and the lips were slightly constricted by girare. The area and length sensitivity functions indicated that all these changes in the vocal tract contributed to decreasing R2. Then, to examine effects of decreasing R2 on the power of 2f0 at G5, the area function whose R1 and R2 were tuned to f0 and 2f0 respectively was modified to decrease only R2 by more than 300 Hz, by the resonance tuning method proposed by Adachi et al. [6]. As a result, the power of 2f0 in the synthesized vowel reduced by approximately 12 dB. Thus, at least for this subject, it is safe to say that girare was a skill which reduced the power of 2f0 by reducing R2 at high pitches where R1 and R2 are normally tuned to f0 and 2f0, respectively.

However, it is unclear whether the skill of this subject revealed in the present study is common among sopranos and can be defined as girare in general. Through informal interviews with several sopranos, we found that the definition of girare was ambiguous and varied among them; this is understandable, as singing skills including girare are basically invisible. Thus, although the physiological mechanism and acoustic function of girare for the subject were revealed in the present study, because this is just a case study, more subjects are necessary for generalization. For that purpose, the direct measurement of the vocal tract resonances and voice harmonics [1, 2] could be more convenient than MRI experiments.

Acknowledgments

This study was partly supported by JSPS KAKENHI Grant Numbers 24500233 and 22520156. The authors thank Dr. Parham Mokhtari for helpful comments in revising the manuscript.

5. REFERENCES