"Ring" in the Solo Child Singing Voice

***David M. Howard, †Jenevora Williams, and ‡,§Christian T. Herbst,** **York, †Guildford, UK, and ‡Olomouc, Czech Republic, and §Vienna, Austria*

Summary: Objectives/Hypothesis. Listeners often describe the voices of solo child singers as being "pure" or "clear"; these terms would suggest that the voice is not only pleasant but also clearly audible. The audibility or clarity could be attributed to the presence of high-frequency partials in the sound: a "brightness" or "ring." This article aims to investigate spectrally the acoustic nature of this ring phenomenon in children's solo voices, and in particular, relating it to their "nonring" production. Additionally, this is set in the context of establishing to what extent, if any, the spectral characteristics of ring are shared with those of the singer's formant cluster associated with professional adult opera singers in the 2.5–3.5 kHz region.

Methods. A group of child solo singers, acknowledged as outstanding by a singing teacher who specializes in teaching professional child singers, were recorded in a major UK concert hall performing *Come unto him, all ye that labour*, from the aria *He shall feed his flock* from The Messiah by GF Handel. Their singing was accompanied by a recording of a piano played through in-ear headphones. Sound pressure recordings were made from well within the critical distance in the hall. The singers were observed to produce notes with and without ring, and these recordings were analyzed in the frequency domain to investigate their spectra.

Results. The results indicate that there is evidence to suggest that ring in child solo singers is carried in two areas of the output spectrum: first in the singer's formant cluster region, centered around 4 kHz, which is more than 1000 Hz higher than what is observed in adults; and second in the region around 7.5–11 kHz where a significant strengthening of harmonic presence is observed. A perceptual test has been carried out demonstrating that 94% of 62 listeners label a synthesized version of the calculated overall average ring spectrum for all subjects as having ring when compared with a synthesized version of the calculated overall average nonring spectrum.

Conclusions. The notion of ring in the child solo voice manifests itself not only with spectral features in common with the projection peak found in adult singers but also in a higher frequency region. It is suggested that the formant cluster at around 4 kHz is the children's equivalent of the singers' formant cluster; the frequency is higher than in the adult, most likely due to the smaller dimensions of the epilaryngeal tube. The frequency cluster observed as a strong peak at about 7.5–11 kHz, when added to the children's singers' formant cluster, may be the key to cueing the notion of ring in the child solo voice.

Key Words: Child voice–Singing–Soloist–Ring–Singer's formant–Singer's formant cluster.

INTRODUCTION

At the start of a Service of Lessons and Carols at Christmas, pioneered by the King's College Cambridge, UK and broadcast worldwide, a lone boy chorister sings the first verse of "*Once in royal David's city.*"[1, p100] Listeners have described children's voices as having "lightness and clarity" and "beauty of tone,"[2] a "clean white tone,"[3] and as being "clearer," "purer," "echoey," and "non fuzzy."[4] Day[5] lists epithets that have been used to characterize children's singing including, "pure," "sweet," "other worldly," "ethereal," and "impersonal." In general, listeners seem able to pick out child solo singers who exhibit a ring-like sound easily and by mutual assent, and the term "ring" is often used in common parlance to describe it. In the solo performance context, such a sound tends to project well in a large reverberant building. The concept of ring is not restricted to solo child singers; it is also heard in the outputs from trained adult singers[6, p47, 7, 8, p55] and Titze[9, p41] describes a "ringing" voice quality as a "brilliant sound—has ping in it."

Because "ring" appears perceptually to have features in common with the projection of the sung sound, it has been acoustically linked[10] to the singer's formant[7,9,11] or singer's formant cluster[12] that is associated with how a singer can be heard above an orchestra[13] and "projecting or focusing the voice."[14, p71] Child singers are also able to project their singing voices; a trained solo voice can be heard clearly in a large building or over accompaniment such as an organ or orchestra. The vocal quality linked with this projected sound is often referred to, in adult's voices, as "ping" or "ring"[3,15] or, in the Bel Canto tradition, "squillo" or "twang."[16]

The adult listener has a response to the solo voice of the child singer that is reflected in many cultures. The reason for this could be linked to the potential for perceived vulnerability in the young singer, which in the performance is combined with an advanced level of artistry. Or it could be that there is an element of the acoustic output of a child's voice that is particularly emotive for the adult listener? Emotional triggers in the singing voice tend to be sounds that have similarities with primal emotive vocal gestures such as crying, wailing, calling, giggling, or sighing.[17–19] Scheiner et al[20] reported a significant increase in the overall frequency range of the vocal output from positive (surprise, interest, contentment, joy) to negative (pain, unease, anger) emotional nonverbal vocalizations (cry, coo,

Accepted for publication September 4, 2013.

From the *Department of Electronics, York Centre for Singing Science, University of York, York, UK; †Singing teaching, Guildford, UK; ‡Department of Biophysics, Palacky University Olomouc, Olomouc, Czech Republic; and the §Department of Cognitive Biology, Bioacoustics Laboratory, University of Vienna, Vienna, Austria.

Address correspondence and reprint requests to David M. Howard, Department of Electronics, Audio Laboratory, University of York, Heslington, York, YO10 5DD, UK. E-mail: david.howard@york.ac.uk

Journal of Voice, Vol. 28, No. 2, pp. 161-169
0892-1997/$36.00

http://dx.doi.org/10.1016/j.jvoice.2013.09.001

wail, moan, babble, whoop, squeal, laugh, hic, groan, croak, raspberry) by infants in their first year of life.

Is the notion of ring when listening to a child solo singer an acoustically measurable phenomenon? A performer can elicit an emotional response in a listener, especially if the listener empathetically "tunes in" to the extreme physicality of the performer's gesture, recognizable by the performer's use of a vocal gesture that is far removed from everyday speaking such as a very high pitch, sustained note or phrase, or projected voice. The listener may also "tune in" to primal emotions expressed in the vocal quality. When a child sings a solo, the emotional response of a listener is typically different to that experienced when an adult sings that same solo. This article seeks to explore the acoustic nature of the ring in the solo singing of children. In particular, it compares the spectra of sung sounds produced by solo child singers in the context of sung outputs that exhibit ring with those that do not have ring.

METHOD

A group of 10 child solo performers (seven boys and three girls) were selected by a professional singing teacher (author J.W.), who works extensively with child singers and is the vocal advisor to the UK National Youth Choir, as being exemplar singers exhibiting ring in their sung output. Their ages ranged from 9 years 10 months to 14 years 4 months; individual ages are reported in Table 1. All were judged to be prepubescent at the time of the recording; the girls had not yet experienced their first menstrual cycle, and the boys had an average speaking fundamental frequency (F_0) of A3 (220 Hz) or higher.

To ensure that the soloists were focused on giving a performance during the recording, the experiment was conducted on the stage of the Menuhin concert hall at the 316-seat Yehudi Menuhin School in Cobham in Surrey, UK.[21] Although there was no audience present during the recordings apart from accompanying persons and authors D.M.H. and J.W., singing on stage in a 300-seat professional concert hall gave each soloist an audio and visual experience that was more commensurate with a performance than is typically found in such studies in which small recording rooms are used. To assess a vocal performance, it was considered important to use a representative performance space rather than use a laboratory recording room.

The task was to perform the section of the aria *He shall feed his flock*, which starts *Come unto him, all ye that labour*, from *Messiah* by GF Handel.[22, p83] A piano accompaniment was provided via an iPod Touch and in-ear earphones (Apple Inc, Cupertino, CA) to provide both a tempo and pitch reference throughout that was audible only to the performer. Thus, the audio recordings were of the vocal output alone, and it is demonstrated below that the recorded sound level was well above the local acoustic noise floor of the concert hall. A professional soprano also took part in the experiment and she followed exactly the same recording protocol in order that comparisons could be made between her acoustic output and those of the child soloists.

It was noted that the greatest perceived ring occurred toward the upper end of the pitch range for all singers. In particular, the word "take" sung on F5 (second F above middle C) at the start of the line *take his yoke upon you* [22, bar 38] was identified by the authors D.M.H. and J.W. as having definitive ring characteristic. The ring quality was heard in the middle portion of the note; the beginning and end of the note tended to have a "nonring" quality. Therefore, after recording the section of the aria, subjects were asked to sing the single word take again, doing their best to exaggerate the difference between maximum and minimum ring on a sustained F5. The singers appeared to understand when they were producing a tone with and without ring. In practice, they were able to enhance this aspect of their singing sound at will.

TABLE 1.
Subject Identification, Gender, Age, and Relative Level of the Maximum Peak (in All Cases, This Was the F_0) Between the "Ring" Version and the "Nonring" Version for Each Subject

ID	Gender	Age	Relative Max. Level (Ring − Nonring) (dB)
F1	Female	11 y 7 mo	5.61
F2	Female	12 y 11 mo	2.94
F3	Female	10 y 0 mo	10.74
M1	Male	13 y 3 mo	2.46
M2	Male	12 y 7 mo	5.28
M3	Male	13 y 7 mo	6.03
M4	Male	12 y 1 mo	6.68
M5	Male	9 y 10 mo	2.38
M6	Male	14 y 4 mo	5.14
M7	Male	14 y 0 mo	8.99
A1	Female	Adult	1.31

The acoustic recording was made using a DPA 4060 omnidirectional microphone (DPA Microphones, Gydevang, Denmark) and a Sennheiser MKH-20 P48 omnidirectional microphone (Sennheiser electronic GmbH & Co., Wedemark, Germany) placed at a distance from the singer's lips of 30 cm at an angle of 45° and 1 m, respectively. The audio data were recorded on two channels of a Sound Devices 744T four-channel digital audio recorder (Sound Devices LLC, Reedsburg, WI,) at a sampling rate of 44.1 kHz and 16-bit resolution to both its internal hard drive and compact flash card simultaneously.

The core analysis in terms of investigating the spectral differences between sung notes with ring and those without ring was a 4096-point Hamming-windowed long-term average spectrum (LTAS) carried out using Wavesurfer (KTH, Stockholm, Sweden). Data processing and graph preparation were carried out using an Excel spreadsheet (Microsoft; Redmond, Washington, DC).

RESULTS

The audio data were transferred digitally to a PC for processing via the compact flash card from the 744T audio recorder. The level of the background and recording system noise was established for the complete recording chain by taking the LTAS of the recorded signal during a period of silence, and this is presented with the first LTAS plots of sung data in Figure 1. It can be seen that the signal level of interest is well above that of the background noise, and because the overall gains of the