The Singing Power Ratio as an Objective Measure of Singing Voice Quality in Untrained Talented and Nontalented Singers

*Christopher Watts, †Kathryn Barnes-Burroughs, ‡Julie Estis, and ‡Debra Blanton

*Harrisonburg, Virginia, †Lubbock, Texas, and ‡Mobile, Alabama

Summary: A growing body of contemporary research has investigated differences between trained and untrained singing voices. However, few studies have separated untrained singers into those who do and do not express abilities related to singing talent, including accurate pitch control and production of a pleasant timbre (voice quality). This investigation studied measures of the singing power ratio (SPR), which is a quantitative measure of the resonant quality of the singing voice. SPR reflects the amplification or suppression in the vocal tract of the harmonics produced by the sound source. This measure was acquired from the voices of untrained talented and nontalented singers as a means to objectively investigate voice quality differences. Measures of SPR were acquired from vocal samples with fast Fourier transform (FFT) power spectra to analyze the amplitude level of the partials in the acoustic spectrum. Long-term average spectra (LTAS) were also analyzed. Results indicated significant differences in SPR between groups, which suggest that vocal tract resonance, and its effect on perceived vocal timbre or quality, may be an important variable related to the perception of singing talent. LTAS confirmed group differences in the tuning of vocal tract harmonics.

Key Words: Singing talent—Singing ability—Resonance—Singing power ratio—Phonatory control.

INTRODUCTION

The facility to express a skill in the performance of a task is often credited to talent. Talent can be thought of as a special natural ability that has the potential to lead to a capacity for achievement or success. Anecdotally, physiological talent is commonly recognized, most notably, in sports performance. Singing talent is also recognized, such as when judges critique the degree of talent expressed by a person for making decisions regarding awards or admittance to specialized training programs. Singing
Singing involves both physiological and artistic components. Physiologically, singing talent can be defined as the special natural ability to coordinate and control the musculature and structures of the voice organ to produce musical modulations of the voice where the sounds vary over a wide range of frequencies and are in tune with each other, or where such sounds are melodic. Singing talent can be inferred from listening to a person sing. In fact, many professional voice teachers have frequent experience making these types of judgments. It is likely that singing talent is expressed along a continuum so that there are different degrees of ability. However, this fact is not necessarily taken into consideration universally in many real-world contexts.

Specifically, binary judgments regarding singing talent (e.g., present vs. not present; at a required level vs. not at a required level) are routinely applied by experienced judges to persons competing for admission to college music programs, for voice scholarships, and in other competitions.

Little empirical data are available that investigate, physiologically, why some people, who are untrained, express singing talent, whereas others do not. Goetze et al. suggested several variables that are requisite for accurate singing (accurate singing, or singing in tune, is one variable thought to be highly important for the expression of singing talent), which included the ability to discriminate between pitches, vocalize over a wide range of pitches, the ability to monitor vocal pitch, and a motivation to sing. The importance of pitch control for accurate singing has also been reported by other authors.

To investigate which perceptual factors were most important for judgments of singing talent, Watts et al. conducted a large survey of professional voice teachers. Over 1000 participants responded to questions that queried what factors were most important in judging whether an untrained voice expressed singing talent. According to these experts, the three most important variables were intonation, timbre, and musicality. The first two variables, intonation and timbre, related to physiological singing abilities (that is, pitch control and voice quality, respectively). Both factors can be measured objectively or experienced perceptually by a listener. Intonation is influenced by pitch control abilities. Timbre, or voice quality, is influenced by the supraglottal configuration of the vocal tract and its resonant effects on the voice source.

It is possible to objectively investigate pitch control and voice quality in talented and nontalented persons via acoustic analyses of vocal productions. Study of these variables may provide evidence toward understanding the physiological capabilities that distinguish one group from the other. The singing power ratio (SPR) has objectively studied singing voice quality in both trained and untrained singers. This measure is an indicator of resonant tuning in the vocal tract, and evidence has suggested that SPR represents the acoustic characteristics of singing voice quality. SPR is calculated by measuring the ratio of the peak intensities between the 2–4-kHz and 0–2-kHz frequency bands in the context of sustained vowels or vocalic segments in sung/spoken samples. Lower SPR measures indicate greater energy in the higher harmonics, a condition that has been said to enhance the richness and perception of “ring” in the voice, and thus it influences the perception of voice quality in a positive way. This measure has been found to differentiate the singing voices of trained and untrained singers, as well as spoken and sung productions.

Judgments of singing talent are usually exclusively subjective. However, perceptual experience of voiced sound is influenced by the spectral relationship of the harmonics in the sound reaching the ear. If judges are asked to include timbre as a factor in rating singing talent, then analysis of the voice spectrum can assist in determining the influence of timbre on those judgments. SPR is one type of measure that we can use toward this goal. Another measure from which voice quality can be inferred is long-term average spectra (LTAS). LTAS measures allow inferences regarding the sound source spectrum and can be calculated from connected speech samples, whereas SPR focuses more on the influence of resonance in the supraglottal vocal tract and its measurement is restricted to the context of vocalic productions. LTAS averages spectral features of the speech signal over time, from which aspects of voice quality (timbre) can be inferred via...
calculation of spectral slope or tilt. Spectral tilt represents the drop in energy from lower frequency harmonics to the higher frequency harmonics in a spectrum. LTAS is averaged over time, and it analyzes longer utterances that include speech. Therefore, LTAS is a useful tool for analysis of voice quality in sung passages.

This study is one of a series of investigations focusing on objective measures of phonatory, perceptual, and resonant tuning abilities in untrained talented and nontalented singers, where perceptual judgments of singing talent are based on intonation accuracy and voice timbre (eg, voice quality). The goal of this line of research is to collect physiological evidence, both indirectly and directly, for the purpose of explaining differences in singing talent in untrained persons. The specific research questions of this study are as follows: (1) Do untrained persons with talented singing voices resonate the source spectrum (eg, modify the energy relationships of harmonics produced by the vibrating vocal folds) differently when singing compared with untrained nontalented persons? (2) Do untrained persons with singing talent demonstrate a different degree of spectral tilt when singing compared with untrained nontalented persons? It is hypothesized that persons judged to express a talented singing voice will manifest lower SPR measures and less spectral tilt, both of which would be indicative of greater energy in the higher harmonics, compared with persons judged to lack singing talent.

**METHOD**

**Participants**

Participants included 39 persons recruited from the undergraduate and graduate populations of speech and hearing majors at the University of South Alabama to participate in an ongoing line of investigation into singing accuracy and talent. Each participant had the potential to be assigned to one of two groups, which, for the purposes of this study, were labeled as the talented and nontalented groups. Group assignment was based on judgments made by professional voice teachers, as described below. Participants were all women, and they ranged in age from 20 to 35 years (mean = 23 years).

Only data from 33 participants were included in the final analysis. Six subjects either did not meet inclusion criteria or were not rated unanimously by judges as falling into one of the experimental subgroups (see below). All participants had normal hearing at 20-dB HL for 500 through 4000 Hz, no history of chronic vocal pathology, no previous voice therapy, no history of drug or alcohol abuse, no history of smoking, and reported no allergies or voice problems at the time of testing. In addition, no subjects had received one-on-one voice training from either a professional or amateur teacher.

To assign participants into experimental groups, recordings of a song stimulus produced by each participant were presented to two trained professional teachers of singing. Both judges were women, possessed graduate degrees in voice performance, were voice faculty in a university setting with experience at rating singing talent, and had over 5 years of professional pedagogical experience. Each judge listened to the recordings independently. The identities of the singers were not disclosed to the judges. The judges were provided the singing samples on a compact disk and asked to judge whether the person singing expressed singing talent. The judges were instructed to make binary judgments (present / not present) of perceived talent, and to base their judgments solely on the intonation (pitch accuracy) and timbre (voice quality) of the voice during singing. Each judge reported understanding of the instructions and indicated that they felt comfortable, and had experience, with those two variables as criteria for rating vocal talent.

Only subjects who were rated unanimously by both judges as either expressing or not expressing singing talent were assigned to one of the two groups. Any subject who was rated differently by the two judges was excluded from the study. Out of the 39 initial participants, judges agreed on 34 ratings (87%). Twelve subjects were rated unanimously as expressing accurate intonation and a voice quality pleasant enough to be characterized as talented, whereas 22 participants were rated unanimously as nontalented. One subject rated as nontalented was found to have a history of neurological problems and excluded from the study. To assess reliability of judgments, a third professional voice
teacher listened to 13 randomly selected (33%) stimuli sung by the pool of participants who were rated unanimously by the two original judges. There was 100% agreement (6 talented; 7 nontalented) between the ratings of the third judge and those of the first two, which suggests that the original ratings of talent/nontalent based on intonation and timbre were reliable.

**Procedures**

Before testing, participants read and signed a consent form, completed questionnaires, and underwent audiometric screening to ensure hearing was within the required limits for the study. Participants were then asked to sing the first stanza of the song “America the Beautiful” in their best singing voice at comfortable pitch and loudness levels. To facilitate the best singing sample from participants, they were given a minimum of 5 minutes for vocal warm-up, including time to practice the stimulus song. Every participant reported that they were familiar with the song. If any participant was unfamiliar with the tasks for vocal warm-up, they were given instruction. Finally, participants were informed that no one could hear them while singing, as they were to be alone in a sound-treated booth during production. Recording took place in a double-walled sound booth with the subject in isolation, and sung samples were recorded digitally at a sampling rate of 44.1 kHz onto an Alesis Masterlink (Alesis West Coast, Los Angeles, CA) digital recorder via a direct line input from an AKG Acoustics (AKG USA, Nashville, TN) head-mounted microphone positioned 3 cm off-center at the left corner of the mouth of the subject. These samples placed each participant into their respective group based on the perceptual judgments of the professional voice teachers (see above), and for later spectral analysis.

**Data calculation and analyses**

Two measures were calculated to compare harmonic energy in the two groups of subjects: SPR and LTAS. For SPR measures, fast Fourier transform (FFT) power spectra were calculated from selected vocalic segments (five total) in the recordings of the stimulus song with the Kay Elemetrics Computerized Speech Lab (CSL, model 4400; Kay Elemetrics Corporation, Lincoln Park, NJ), with a time window of 1024 points and a bandwidth of 8000 Hz. For each vocalic segment, energy peaks were calculated between 0 and 4000 Hz. SPR was then calculated by subtracting the amplitude of the strongest partial between 2 and 4 kHz from the level of the strongest partial between 0 and 2 kHz. This measure has been reported to reflect the “ring” in the voice, and it corresponds to the resonant quality of the singing voice.\(^9,10\)

LTAS were also calculated by analyzing the stimulus songs via the CSL. LTAS were calculated in this study as a measure of spectral tilt of the frequencies between 0 and 8000 Hz, in increments of 500 Hz, for comparison with the SPR measures. LTAS were calculated from the entire duration of each song sung by the subjects, and the harmonic peaks were averaged across each group for intergroup comparison. Because loudness variations can affect the intensity in higher spectrum partials, the amplitude values of the spectral peaks were normalized to control for loudness variations between subjects. This process was accomplished by assigning the intensity of the strongest partial a value of 1 and each subsequent partial a proportional value compared with this peak intensity.

An independent samples \(t\) test was applied to the SPR data to investigate differences between the two experimental groups, with the significance level set at \(P \leq 0.05\). Averaged LTAS data were examined graphically to assess spectral profiles in the two groups, and compared with the SPR results.

**RESULTS**

Preliminary comparisons of subject performance revealed apparent dissimilarities in performance between the two groups. Table 1 displays measures of SPR in subjects of both groups. Collectively, the talented group manifested an SPR that was approximately 8 dB less than the nontalented group. Almost half of the nontalented subjects manifested SPR values over 30 dB, whereas only one talented subject produced a value that high. The variance was very similar (6.9 dB for talented, 5.7 for nontalented) between the two experimental groups. The data warranted more investigation to test for group differences.
TABLE 1. Individual Means and Group Means With Standard Deviations on Measurements of Formant Power Difference (SPR) in dBSPL

<table>
<thead>
<tr>
<th>Subject</th>
<th>Talented</th>
<th>Nontalented</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>19.3</td>
<td>27.8</td>
</tr>
<tr>
<td>2</td>
<td>20.6</td>
<td>24.0</td>
</tr>
<tr>
<td>3</td>
<td>18.9</td>
<td>31.1</td>
</tr>
<tr>
<td>4</td>
<td>22.0</td>
<td>29.4</td>
</tr>
<tr>
<td>5</td>
<td>17.9</td>
<td>26.3</td>
</tr>
<tr>
<td>6</td>
<td>18.9</td>
<td>29.5</td>
</tr>
<tr>
<td>7</td>
<td>34.1</td>
<td>33.6</td>
</tr>
<tr>
<td>8</td>
<td>24.6</td>
<td>26.6</td>
</tr>
<tr>
<td>9</td>
<td>26.5</td>
<td>28.6</td>
</tr>
<tr>
<td>10</td>
<td>18.8</td>
<td>31.8</td>
</tr>
<tr>
<td>11</td>
<td>21.2</td>
<td>33.6</td>
</tr>
<tr>
<td>12</td>
<td>28.5</td>
<td>36.0</td>
</tr>
<tr>
<td>13</td>
<td>–</td>
<td>29.9</td>
</tr>
<tr>
<td>14</td>
<td>–</td>
<td>33.6</td>
</tr>
<tr>
<td>15</td>
<td>–</td>
<td>31.5</td>
</tr>
<tr>
<td>16</td>
<td>–</td>
<td>32.8</td>
</tr>
<tr>
<td>17</td>
<td>–</td>
<td>30.9</td>
</tr>
<tr>
<td>18</td>
<td>–</td>
<td>28.5</td>
</tr>
<tr>
<td>19</td>
<td>–</td>
<td>28.6</td>
</tr>
<tr>
<td>20</td>
<td>–</td>
<td>34.7</td>
</tr>
<tr>
<td>21</td>
<td>–</td>
<td>37.0</td>
</tr>
<tr>
<td>Group avg.</td>
<td>22.6 (6.9)</td>
<td>30.7 (5.7)</td>
</tr>
</tbody>
</table>

**Figure 1** illustrates a box plot representing averaged SPR values (in dBSPL) for subjects in both experimental groups. Again, higher values of SPR represented less energy in the higher harmonics (above 2000 Hz) in ratio to the energy in the lower harmonics (below 2000 Hz). Although the variability of the talented group seems to be high, the superior vertical line on the talented plot represents only one subject in this group who produced an SPR value that was more than one standard deviation from the mean (talented subject 7). Conversely, only one subject in the nontalented group (nontalented subject 2) produced a SPR value less than one standard deviation from the mean, as illustrated by the inferior vertical line on the nontalented plot. Collectively, the separation in means between the two groups is apparent.

The Levene test for equality of variances was above 0.05, so the data were analyzed assuming equal group variances. Results from the *t* test revealed a significant difference between the two groups with regard to SPR (*t* = 8.13, *P* < 0.001; 95% CI = 9.92–6.04). SPR scores were significantly lower for the talented group compared with the nontalented group. Lower SPR scores correspond to an increase in perceived vocal “ring” and the perception of enhanced voice quality (timbre) because of the greater energy in the higher harmonics. The level of significance and confidence interval suggested a robust finding.

**Figure 2** illustrates averaged LTAS results for each group. As can be seen, the talented subjects displayed lower spectral tilt, which means that the higher harmonics were more dominant in the spectrum compared with the nontalented group, where the fundamental and lower harmonics were more dominant. The talented group demonstrated a noticeable rise in energy throughout the 2500-Hz to 4000-Hz frequency band. These measures agreed with the results from the SPR data.

**DISCUSSION**

This study compared measures of the SPR, which reflected harmonic tuning of the source spectrum by the superior vocal tract, in untrained talented and
nontalented singers. We used a relatively small sample size for group comparisons, and as such any generalizations based on the results should be made with caution. As expected, a significant difference was found between the two groups. The talented subjects were collectively different from the nontalented subjects with regard to how higher spectral partials were tuned. There was greater emphasis in the 2–4-kHz harmonic band, which suggests more “ring” to the singing voice. Greater energy in this spectral region has been argued to reflect on singing voice quality.\(^{13}\) If the SPR is a measure of singing voice quality, it may be possible to infer that the energy relationship between the lower and the higher harmonics was a perceptually meaningful factor to the judges who placed each untrained person into their respective group. Specifically, those with greater energy in the higher harmonics (corresponding to a lower spectral tilt) were judged as expressing singing talent.

The results of this study are consistent with previous investigations that have studied harmonic tuning measures (e.g., the formant of the singer) in trained and untrained singers. These studies have typically reported greater energy in the region of 2–4 kHz in trained voices.\(^{9,10,14}\) The results of this study are also in line with a body of research that has demonstrated that trained singers, in comparison with nonsingers, exhibit different physiological capabilities during the act of singing.\(^{14–21}\) Measures of harmonic tuning, such as SPR, offer an indirect method of investigating physiological abilities during singing. Harmonic tuning of the vocal tract is partially dependent on the given anatomical structure of a person, but also dependent on the physiological ability to manipulate the vocal tract to tune the voice spectrum for the production of a target voice quality.

Judgment related to the presence or absence of expressed singing talent is also influenced by pitch control.\(^{8}\) Previous studies have found that subjects judged to produce talented singing voices show more finely tuned pitch perception capacities and are more accurate at vocal pitch-matching tasks.\(^{19}\) Combined with findings from this study, the data support the views of many singing pedagogues that pitch control and timbre are the two most salient aspects of the singing voice for perceptual judgments of talent in untrained singers.\(^{8}\)

Hollien\(^{22}\) has suggested that superior singing abilities are a product of both talent and training. In the literature of voice science where physiological and perceptual variables related to the singing voice are investigated, persons with and without voice training are often compared. These subgroups are usually labeled as trained and untrained singers, singers and nonsingers, or professional singers and nonsingers. Within a population of untrained singers, some will express singing talent, whereas others will not. Although numerous studies have found differences between persons with and without training, few studies have separated untrained subjects into those who have and do not have expressed singing talent. In addition, although there is great public interest in singing talent, little objective empirical evidence sheds light on the factors responsible for expressed singing talent in untrained persons. This study and previous studies in this line of investigation have presented data that support an argument for more research in this topic, as numerous physiological and perceptual differences may exist, yet they are poorly defined, between untrained talented and nontalented singers.

To date, the line of investigation pertaining to this study has found several interesting results. It seems that professional teachers of singing rely more on pitch control and voice quality than on other factors when making judgments of singing talent.\(^{8}\) When testing actual abilities of talented and nontalented persons, talented singers are more accurate.
than nontalented singers when matching pitch, which is likely related to greater pitch control abilities, including perceptual acuity. The results of this study also delineate differences between these two groups during the act of singing. Specifically, talented singers seem to produce lower spectral tilt when resonating the voice source, which most likely contributes to the perception of “ring” in the voice or, at a minimum, a more pleasing vocal timbre, and possibly increases the likelihood for being perceived as expressing singing talent.

Although this study involved a small sample size, the significant findings point to the need for more in-depth and comprehensive studies that investigate physiological and perceptual differences in talented and nontalented persons to understand why some, without training, can use their vocal mechanism to produce a talented singing voice and others are not. This knowledge is important for the area of voice science, as it will add to our basic understanding of the voice organ and the physiological bases that are most crucial for controlling the voice during the act of singing.

REFERENCES