Soul and Musical Theater: A Comparison of Two Vocal Styles

*Hanna Hallqvist, †,‡Filipa M.B. Lã, and *,§Johan Sundberg, *§*Stockholm, Sweden*, †, *and* ‡*Coimbra, Portugal*

Summary: The phonatory and resonatory characteristics of nonclassical styles of singing have been rarely analyzed in voice research. Six professional singers volunteered to sing excerpts from two songs pertaining to the musical theater and to the soul styles of singing. Voice source parameters and formant frequencies were analyzed by inverse filtering tones, sung at the same fundamental frequencies in both excerpts. As compared with musical theater, the soul style was characterized by significantly higher subglottal pressure and maximum flow declination rate. Yet sound pressure level was lower, suggesting higher glottal resistance. The differences would be the effects of firmer glottal adduction and a greater frequency separation between the first formant and its closest spectrum partial in soul than in musical theater.

Key Words: voice source–inverse filtering–formant frequencies–nonclassical styles–long-term-average spectrum.

INTRODUCTION

Singing in popular music styles is quite widespread, particularly among school children. These styles often raise high demands on vocal endurance and vocal technique, which young singers often lack.[1,2] Therefore, it seems relevant to analyze the phonatory and resonatory characteristics of popular music styles. Such an analysis should contribute to establishing a scientific basis for teachers of singing who guide young voices.

In the last decades, several investigations have been focusing on voice characteristics in nonclassical genres. A great variety of measures have been used, resulting in quite complex descriptions of the phonatory characteristics of styles of singing. For example, Butte and associates analyzed aperiodicity aspects of 1-second long samples taken from recordings of 26 songs that represented six vocal styles.[3] They found statistically significant differences in jitter, shimmer, signal-to-noise ratio, and correlation dimension. Björkner analyzed voice source properties and formant frequencies in male singers performing in the classical operatic and in the musical theater styles[4] and found significant differences with respect to both voice source and formants. The opera singers had lower closed quotient, stronger voice source fundamental, and lower formant frequencies than the musical theater singers.

Attempts have also been made to condense the description of different styles of singing. Thurmer launched the tessiturogram, which showed the pitch content of a song in terms of a histogram.[5] Coleman developed this idea in terms of plots showing sound pressure level (SPL) as function of percent of the total pitch range of the song.[6] Lycke and associates applied the method to 206 female conservatory student singers and found that it clearly separated the voices into three basic types, thus suggesting potential value of the method as a basis for voice classification.[7]

As suggested above, some vocal styles seem associated with specific voice source characteristics, which in turn are typically strongly affected by subglottal pressure (P_{Sub}). Thalén and Sundberg attempted to describe the characteristic combination of this pressure and a phonatory characteristic in terms of a *phonation map*, showing a measure reflecting phonation type as function of P_{sub} (henceforth P_{Sub}).[8] Clear differences between the analyzed styles, blues, pop, jazz, and classical, were found with respect to P_{Sub} and degree of phonatory pressedness.

A phonation map was also used in a subsequent study of voice properties of a professional male singer who sang music excerpts in different styles: rock, soul, pop, and Swedish dance band. Characteristic differences were found with respect to mean P_{Sub}, mean F0, and the average of the normalized amplitude quotient (henceforth NAQ, defined as the ratio between the flow glottogram peak-to-peak pulse amplitude and the product of the period and the maximum flow declination rate, henceforth MFDR).[9] For instance, rock showed high P_{Sub}, high F0, and high degree of glottal adduction, whereas Swedish dance band showed low values in these parameters. Thus, these two styles assumed opposite locations in the phonation map, whereas soul and pop assumed intermediate positions.

In addition to phonatory characteristics, formant frequencies constitute a relevant property of many vocal styles. For example, in a recent investigation comparing belt style and a neutral, nonbelt style, one out of six professional belt singers was found to consistently tune her first formant to a harmonic partial. This of course added to the sound pressure level of the tones.[10] A similar technique has been found also in some folkloristic styles of singing, like Bulgarian female's singing style and in the Persian singing style called Avaz.[11,12]

The purpose of the present study was to compare two styles of singing in the nonclassical repertoire. Soul and musical theater were chosen as these styles appear to be reasonably distinct and can often be professionally performed by the same singer. Our analysis included both voice source and formant frequency characteristics.

METHODS

Six female singers, with an age range of 22–30 years, volunteered as subjects. They had all studied singing for more than

Accepted for publication May 31, 2016.

From the *University College of Music Education, Stockholm, Sweden; †Centre for Social Sciences, University of Coimbra; ‡Institute of Interdisciplinary Research, University of Coimbra, Portugal; and the §Department of Speech Music Hearing, School of Computer Science and Communication, KTH, Stockholm, Sweden.

Address correspondence and reprint requests to Johan Sundberg, Department of Speech Music Hearing, School of Computer Science and Communication, KTH, Stockholm, Sweden and University College of Music Education, Stockholm, Sweden. E-mail: pjohan@speech.kth.se

Journal of Voice, Vol. ■■, No. ■■, pp. ■■-■■
0892-1997

http://dx.doi.org/10.1016/j.jvoice.2016.05.020

FIGURE 1. Music excerpts used in the experiment representing musical theater and soul (left and right scores, respectively). The tones analyzed are marked with *circles*.

7 years, and had considerable experience performing both in soul and musical theater genres. They all gave their informed consent to participation.

The singers were asked to sing about 60-second long excerpts from two songs, both written for female voice. One was from the Memphis soul repertoire (*Son of a Preacher Man*, lyrics and music by John Hurley and Ronnie Wilkins) and the other from the musical theater repertoire (*On my Own* from *Les Misérables* by Claude-Michel Schönberg) (Figure 1). The total pitch range of both excerpts was similar, G3–A4 in the former excerpt and A3–B4 in the latter excerpt. Both corresponded to a comfortable range for the singers. Also, the distribution of pitches in these two excerpts was similar. The singers sang the examples first with the original lyrics and then replaced each syllable in the lyrics with the syllable /pae/.

The songs were recorded in a sound treated studio ($3 \times 4 \times 2.5$ m). Using a Soundswell workstation (Elektronix NG AB, Täby, Sweden), three tracks were recorded. One track recorded the audio signal, picked up by a head-mounted omnidirectional electret microphone (DPA 4065, Thomann GmbH, Burgebrach, Germany) at a measured distance from the mouth. The microphone signal was amplified by the Symetrix SX 202 Dual Mic Preamplifier (Symetrix Inc., Lynnwood, WA). Sound level calibration was made by recording a vowel sound of constant intensity, the sound pressure level of which was measured at the recording microphone by the Ono Sokki Sound Level Meter LA-210 (Ono Sokki, Yokohama, Japan). On a second track was recorded an electroglottograph signal from the Glottal Enterprises MC 2-1 Two Channel Electroglottograph (Syracuse, NY). Oral pressure was picked up by the pressure transducer contained in a Glottal Enterprises MSIF-2. The transducer was attached to a thin plastic tube, ID = 5 mm, which the subjects held in the corner of the mouth such that it captured the oral pressure. The pressure signal was calibrated by recording pressures, measured by means of a custom-made manometer. Subglottal pressure was measured as the oral pressure during the /p/ occlusion.[13]

The long-term average spectra (henceforth LTAS), bandwidth of 400 Hz, of the examples sung with the original lyrics were obtained from the spectrum section option of the Soundswell workstation and equivalent sound level (henceforth Leq) from the histogram module. Formant frequencies and flow glottogram data were obtained from the custom-made *DECAP* inverse filtering software (Svante Granqvist, KTH). In this software, frequencies and bandwidths are set manually. The software calculates the transfer function associated with the given combination of formant frequencies and bandwidths. The resulting flow glottogram and spectrum, representing the waveform and spectrum of the transglottal airflow, are displayed in quasi-real-time. The program also displayed the derivative of the electroglottograph signal (dEGG), with a time delay set to correspond to the delay of the acoustic signal relative to the EGG.

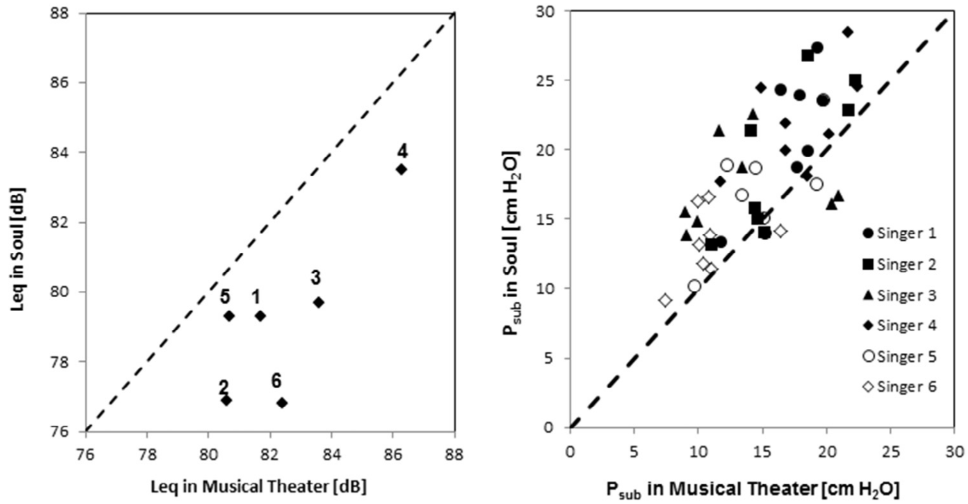

FIGURE 2. Left panel: Equivalent sound level measured in the soul and in the musical theater excerpts. Numbers refer to singers. Right panel: Subglottal pressure used for identical pitches occurring in the soul and in the musical theater excerpt.

In the analysis, formant frequencies and bandwidths were adjusted according to three criteria: (1) ripple free closed phase; (2) voice source spectrum envelope as void of peaks and valleys near formants as possible; and (3) synchrony between the negative dEGG peak and the maximum declination rate of transglottal flow during closure. The resulting flow glottograms were analyzed by means of the custom-made *SNAQ* software (Svante Granqvist). When period and closed phase have been manually marked in the waveform, they yield, in a separated file, fundamental frequency (F0), MFDR, NAQ, dominance of the fundamental, defined as the level difference between the first and the second partial of the source spectrum (H1 – H2), which is defined as the level difference between the first and the second partial of the source spectrum, and closed quotient, defined as the duration ratio between the closed phase and the period (Q_{Closed}).

RESULTS

The singers performed the excerpt in the musical theater style with significantly higher Leq than in the soul style, as illustrated in the left panel of Figure 2. The mean difference in Leq amounted to 3.7 dB. This seems to support the assumption that the singers used higher P_{Sub} in the musical theater style. The right panel of the same figure compares the P_{Sub} used for the same pitches in the two styles. To assess whether there were significant P_{Sub} differences for the tones sung in the musical theater and in the soul styles, a nonparametric paired sample Wilcoxon test was performed. This particular statistical test was used because data showed a skewed distribution. Pressure was significantly higher in the soul style ($z = -2.201$; $P = 0.028$) (Figure 3). On average, across singers and tones, the pressure in musical theater was 0.84 of the pressure in the soul style.

The soul excerpt contained a greater number of short tones than the musical theater excerpt; the average tone duration was 345 ms in soul and 670 ms in musical theater (Figure 1). However, the Leq measure is rather insensitive to pauses in the signal. Eliminating the pauses between tones in the soul excerpt reduced Leq by no more than 1.5 dB. Thus, the reason for the higher P_{Sub} in the soul example did not seem to be the difference in musical structure between the examples.

High subglottal pressures are typically associated with less steep spectrum envelope slopes than low subglottal pressures. The spectrum slopes can be compared in Figure 4. The LTAS curves for musical theater showed a peak near 1000 Hz for all singers except singer 1. This peak was produced by the second partial of the long tone B4, which did not occur in the soul excerpt. Hence, it would not belong to the characteristics of the musical theater style. With regard to frequencies above 2000 Hz, four of the six singers showed essentially the same LTAS curve for both music genres. Thus, the higher P_{Sub} in soul did not produce higher LTAS levels in the high frequency range.

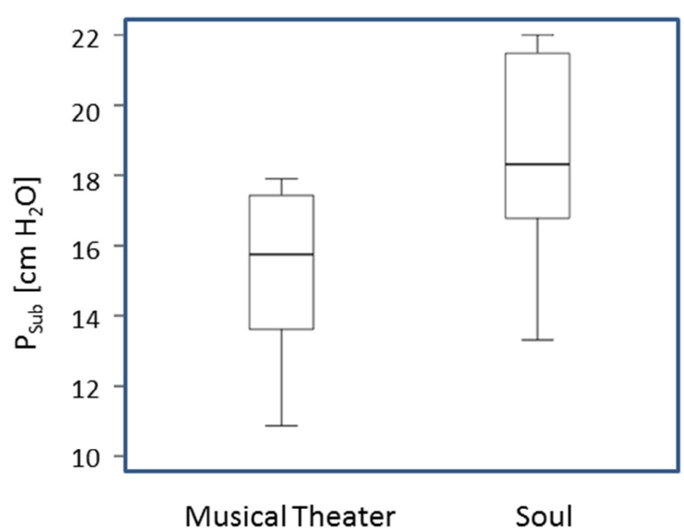

FIGURE 3. Box plot of the subglottal pressure (P_{Sub}) used in the musical theater and soul excerpts. The *boxes* represent the part of the distributions that fall between the 25th and 75th percentiles, the *horizontal lines* represent the medians, and the *vertical lines* connect the lowest and highest values that are not extremes or outliers.

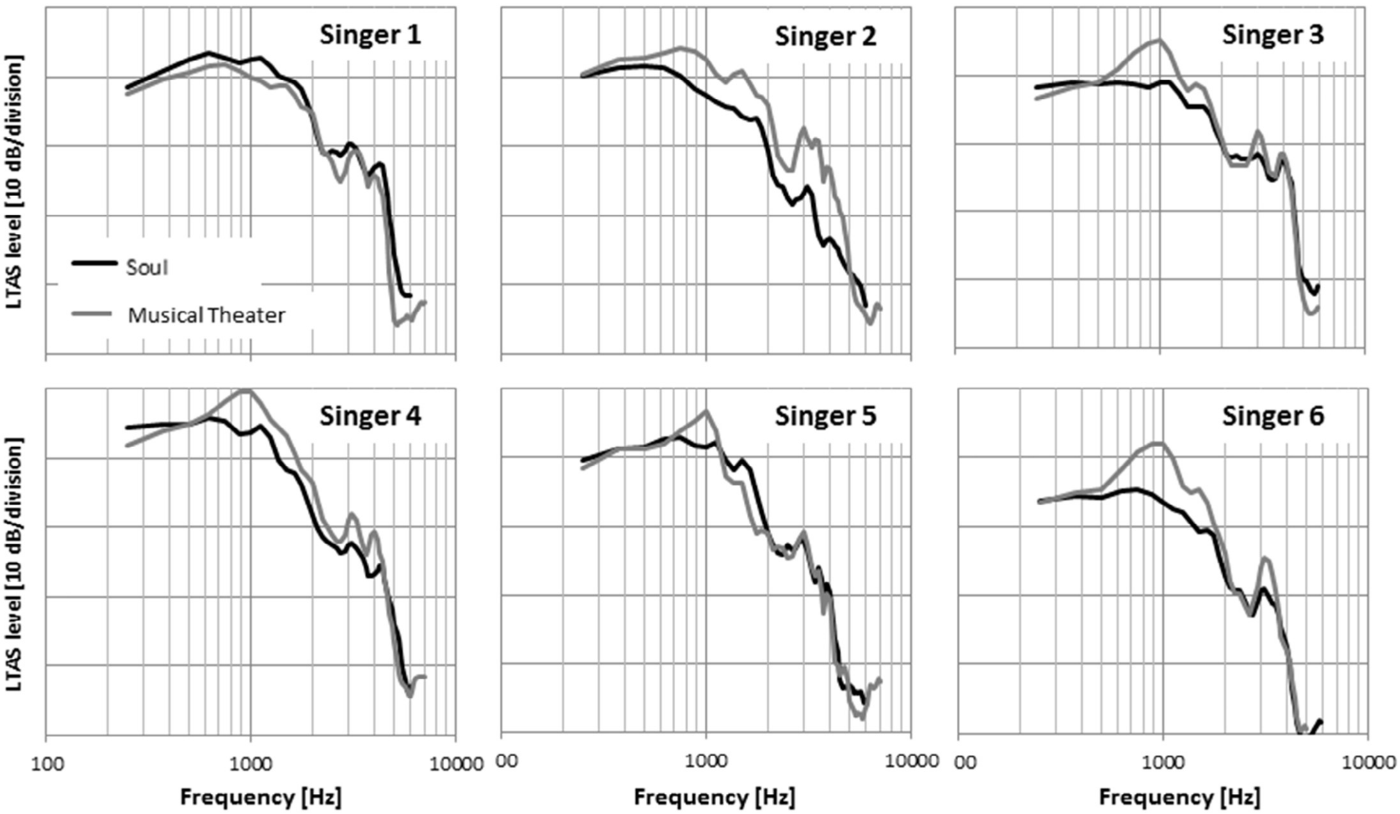

FIGURE 4. Long-term average spectra of the musical theater and the soul excerpts (*gray* and *black curves*, respectively).

Figure 5 compares flow glottogram parameters in terms of the pulse amplitude, MFDR, Q_{Closed}, H1 – H2, and NAQ, which the singers used for identical pitches in the two excerpts. Most singers showed a trend to sing with greater pulse amplitudes, higher MFDR values, and lower NAQ values in soul than in musical theater. This supports the assumption that the singers performed the soul example with firmer glottal adduction, ie, with a somewhat more pressed type of phonation. However, a statistical test failed to support this assumption; a nonparametric paired sample Wilcoxon test only revealed significantly higher MFDR values in soul ($z = -2.201$; $P = 0.028$) (Figure 6).

Figure 7 shows the three lowest formant frequencies used in the inverse filtering analysis for singer 3. In musical theater, this singer tended to tune both F1 and F2 to the vicinity of a spectrum partial. This tendency was observed also for most of the other singers, as illustrated in Figure 8. In the figure, the frequency separation between the formant and its closest partial is expressed in the following way:

$$\Delta f / Fn$$

where Δf is the frequency difference between partial and the formant, and Fn is the frequency of the formant. The columns represent this ratio for F1 and F2, averaged across the inverse filtered tones (left and right panels, respectively). Except for singer 5, the ratios for soul are higher than those for musical theater, indicating that mostly the formants were further away from a partial in the soul style. The effect was weaker for F2 than for F1. However, neither of these effects reached statistical significance.

The P_{Sub}, NAQ, and MFDR results are summarized in Figure 9 in terms of phonation maps, showing averages across pitches and singers. In the graph, the axes of the ellipses represent ±1 standard deviation. Even though the overlap between the ellipses is substantial, it is clear that soul tended to be produced with slightly higher subglottal pressure, higher MFDR, and more glottal adduction than musical theater.

DISCUSSION

This study has shown that the six professional singers used higher P_{Sub} and yet produced lower Leq in soul than in musical theater. This suggests that they used stronger glottal adduction in soul. The trend of the NAQ values to be lower in soul supports the same assumption, even though this trend failed to reach significance.

On the other hand, MFDR, which represents the strength of the vocal tract excitation, was higher in soul. Therefore, soul could be expected to have a higher Leq, but the opposite was found. The reason for these apparently conflicting results may be related to the formant data, which showed that the singers tended to tune F1 closer to a partial, ie, applied more formant tuning, in musical theater. This should increase Leq. Formant tuning has been found to be applied in the belt style of singing,[11] where loud voice is an important characteristic, and also in some folkloristic vocal styles.[12,13]

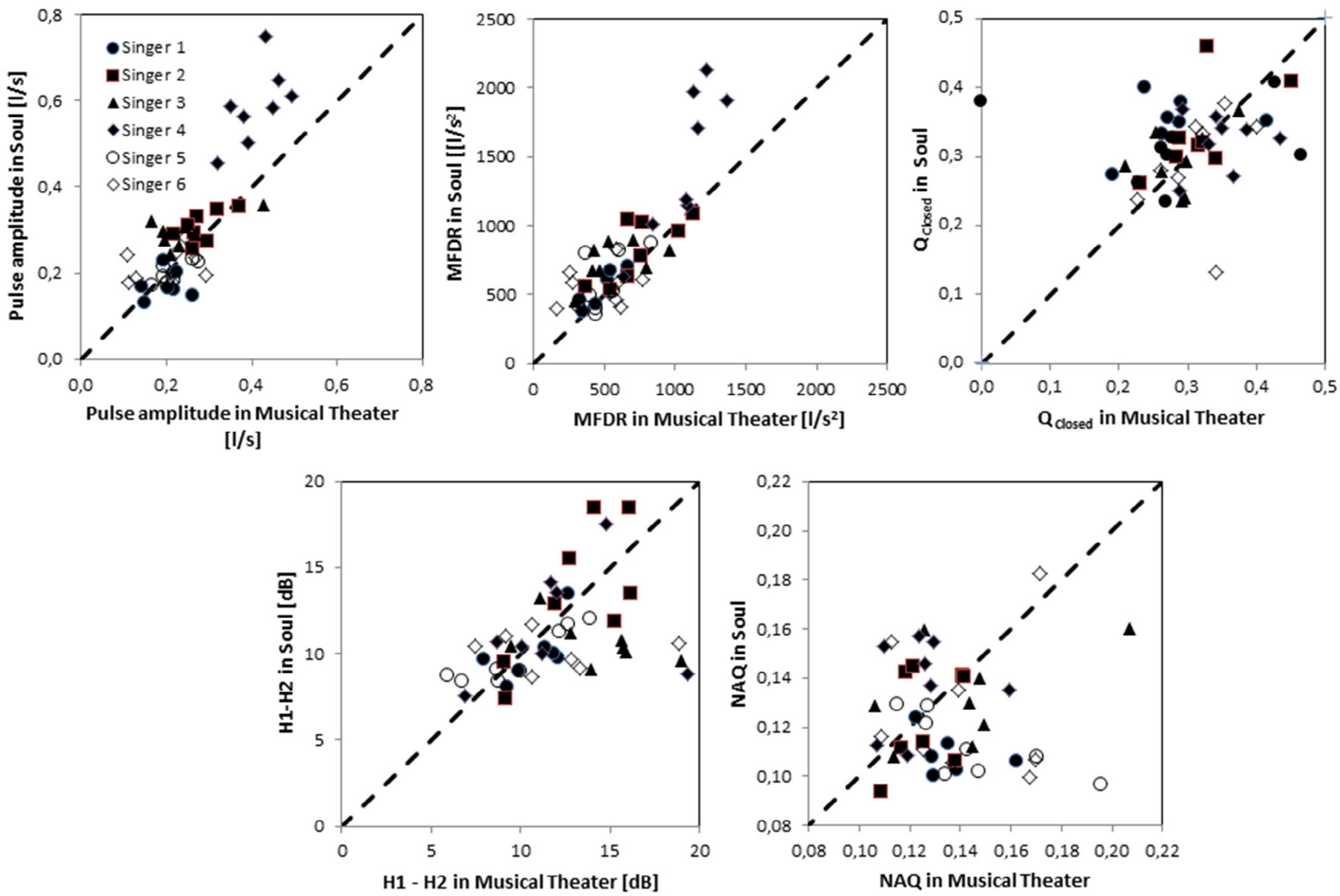

FIGURE 5. Values of pulse amplitude, MFDR, Q_{Closed}, H1 – H2, and NAQ used by the singers for identical pitches in the two styles. MFDR, maximum flow declination rate; NAQ, normalized amplitude quotient.

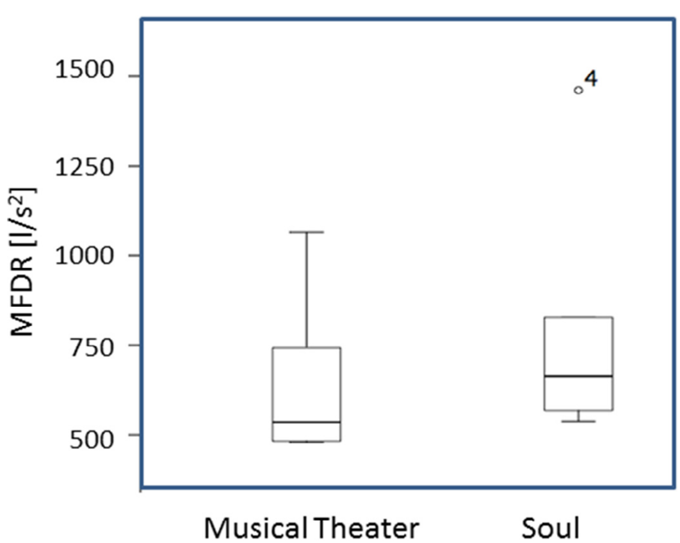

FIGURE 6. Box plot of the MFDR used in the musical theater and soul excerpts. For explanation, see caption of Figure 3. The *circle* represents an outlier value for singer 4. MFDR, maximum flow declination rate.

The differences between the phonatory properties of the two styles were rather small, with respect to NAQ and MFDR, as illustrated by the phonation maps in Figure 9. Much larger differences were found in the investigation where rock, pop, soul, and Swedish dance band were compared in one single singer subject.[9] Rock and Swedish dance band are phonatorily more extreme than the musical theater and soul styles studied in the present investigation. Also, the present study analyzed six singers, and this would have contributed to the scatter of the data.

The question to what extent the singers produced typical examples is important. This question could have been tested experimentally by means of a listening test with an expert panel. Given the recorded material and the number of singer subjects, this did not appear as necessary. All singers worked professionally as soloists in both styles. This suggests that they were skilled and representative in performing in both styles.

Another factor that may have influenced the results is the choice of songs. We choose songs typical of the respective styles, so as to establish experimental conditions that were as musically realistic as possible. The examples had similar pitch ranges. Yet, for each of the pitches analyzed, the musical

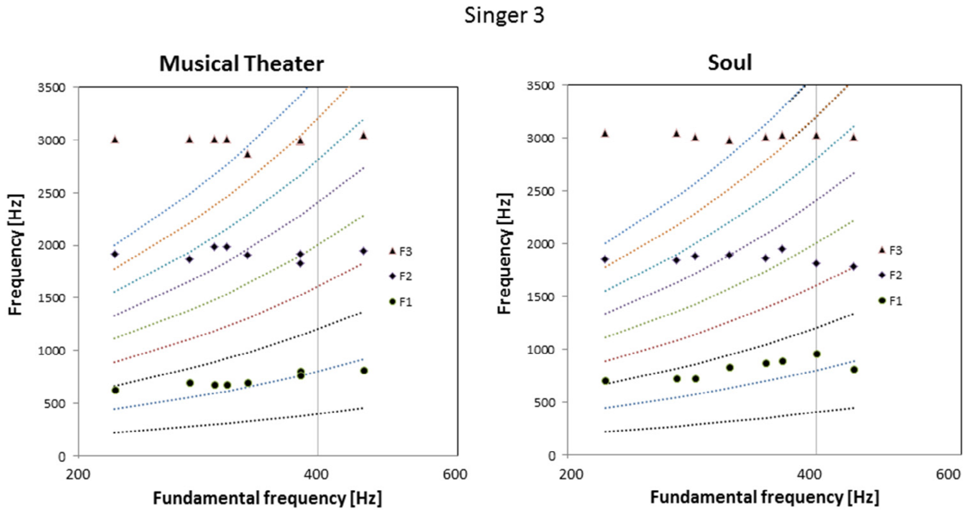

FIGURE 7. F1, F2, and F3 of the vowel /ae/ observed when singer 3 sang the musical theater and the soul examples. The *thin dotted curves* show the frequencies of the spectrum partials.

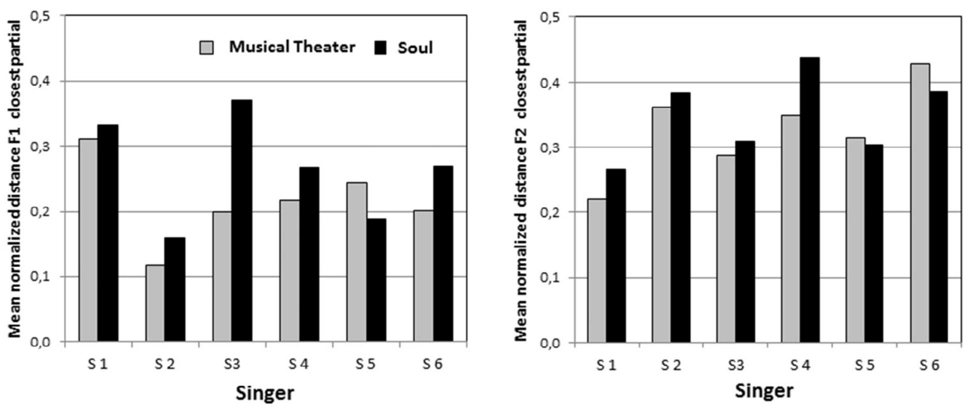

FIGURE 8. Left panel: Average normalized distance between F1 and its closest partial in the indicated singers' performances of the musical theater and soul examples. *Gray* and *black columns* refer to the musical theater and the soul examples. The right panel shows the corresponding values for F2.

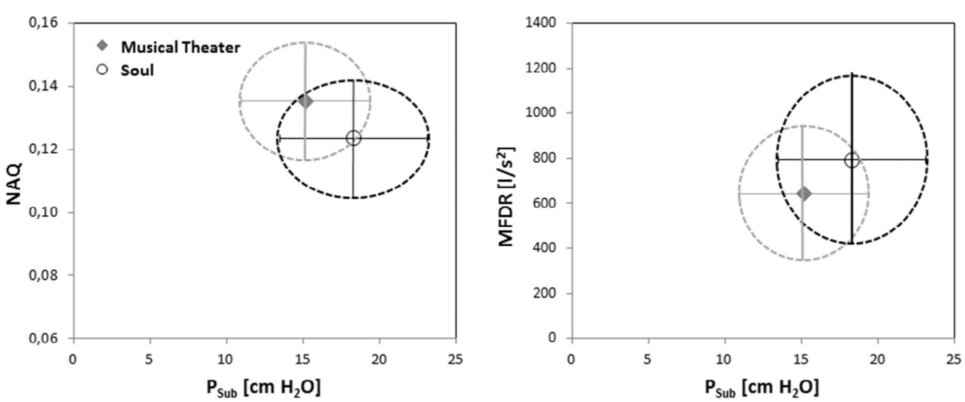

FIGURE 9. Left and right panels show NAQ and MFDR, averaged across subjects and tones, as functions of subglottal pressure for musical theater and soul (*diamond* and *circle*). The gray and black ellipses reflect the standard deviations for the styles, respectively. MFDR, maximum flow declination rate; NAQ, normalized amplitude quotient.

context obviously differed in the two melodies. This may have contributed to the scatter of the data. Asking the singers to sing the same song in both styles appeared to expose the study to the risk of making the subjects confused, possibly producing less typical results. Another experimental design would have been to ask participants to sing scales in the two styles, such that the musical context would be identical. However, the experiment would then be musically less realistic, which was likely to cause a great data scatter.

This investigation focused on the phonatory and resonatory characteristics of musical theater and soul style. However, there would be several other characteristics that belong to the definition of these two styles. Timing and melodic ornaments may be equally important characteristics. For example, riffing is an improvised melodic ornament frequently used in the soul repertoire but rarely in musical theater.

Even though our material was rather small, it showed statistical significance for P_{Sub} which was higher and Leq which was lower in soul than in musical theater. The ratio between these two parameters is related to vocal resistance, which thus was higher in soul. This may be related to vocal register; it seems likely that soul is typically sung in a register with a more modal character than what is commonly used in musical theater singing. In any event, the greater vocal resistance in soul suggests that more care is indicated when teaching young singers to sing soul than to sing musical theater repertoire.

No attempt was made to document a velopharyngeal opening in the singers, eg, in terms of recording nasal flow. However, the voices did not sound nasal. Also, according to Gobl and Mashie in the *Journal of Voice*, the effect on flow glottogram data of a narrow velopharyngeal opening is quite small.[14]

By applying inverse filtering, we could analyze voice source as well as formant frequencies. In voice research, the combination of these factors is commonly analyzed in terms of the radiated spectrum. We would argue that there are good reasons to analyze voice source and formants separately. They are controlled by separate physiological systems—(1) the combination of subglottal pressure and laryngeal adjustments and (2) articulation—so the characteristics of a vocal style may originate from either of these systems or from both. Moreover, meaningful voice source analysis requires measurement also of subglottal pressure as this pressure has a very strong influence on voice source properties.[15,16] Current software for inverse filtering, which returns both flow glottogram parameters and formant frequencies and bandwidths, offers greatly improved possibilities for future analyses of vocal styles.

CONCLUSIONS

Our study has shown that, as compared with musical theater, the soul style of singing was characterized by significantly greater P_{sub} and MFDR and lower Leq. This can be interpreted as a support for the assumption that soul as compared with musical theater was sung in a somewhat heavier register and with greater vocal resistance. This suggests that soul was produced with a firmer glottal adduction, and hence would be vocally more demanding than musical theater. In addition, the frequency distance between the first formant and its closest partial tended to be narrower in musical theater, which should have contributed to a higher Leq.

REFERENCES

1. Guzman M, Lanas A, Olavarria C, et al. Laryngoscopic and spectral analysis of laryngeal and pharyngeal configuration in non-classical singing styles. *J Voice*. 2014;29:130, e21–e28.
2. Barlow C, Lovetri J. Closed quotient and spectral measures of female adolescent singers in different singing styles. *J Voice*. 2009;24:314–318.
3. Butte CJ, Zhang Y, Song H, et al. Perturbation and nonlinear dynamic analysis of different singing styles. *J Voice*. 2008;23:647–652.
4. Björkner E. Musical theater and opera singing—why so different? A study of subglottal pressure, voice source, and formant frequency characteristics. *J Voice*. 2007;22:533–540.
5. Thurmer S. The tessiturogram. *J Voice*. 1988;2:327–329.
6. Coleman RE. Performance demands and the performer's vocal capabilities. *J Voice*. 1987;1:209–216.
7. Lycke H, Decoster W, Ivanova A, et al. Discrimination of three basic female voice types in female singing students by voice range profile-derived parameters. *Folia Phoniatr Logop*. 2012;64:80–86.
8. Thalén M, Sundberg J. Describing different styles of singing: a comparison of a female singer's voice source in "classical," "pop," "jazz" and "blues". *Logoped Phoniatr Vocol*. 2001;26:82–93.
9. Zangger Borch D, Sundberg J. Some phonatory and resonatory characteristics of the rock, pop, soul, and Swedish dance band styles of singing. *J Voice*. 2011;25:532–537.
10. Sundberg J, Thalén M. Respiratory and acoustical differences between belt and neutral style of singing. *J Voice*. 2015;29:418–425.
11. Henrich N, Kiek M, Smith J, et al. Resonance strategies used in Bulgarian women's singing style: a pilot study. *Logoped Phoniatr Vocol*. 2007;32:171–177.
12. Biglari HM, Sundberg J. Timbral and melodic characteristics of Persian and Kurdish singing. *J Acoust Soc Am*. 2012;131:3377.
13. Löfqvist A, Carlborg B, Kitzing P. Initial validation of an indirect measure of subglottal pressure during vowels. *J Acoust Soc Am*. 1982;72:633–635.
14. Gobl C, Mahshie J. Inverse filtering of nasalized vowels using synthesized speech. *J Voice*. 2013;27:155–169.
15. Sundberg J, Andersson M, Hultqvist C. Effects of subglottal pressure variation on professional baritone singers' voice sources. *J Acoust Soc Am*. 1999;105:1965–1971.
16. Sundberg J, Fahlstedt E, Morell A. Effects on the glottal voice source of vocal loudness variation in untrained female and male subjects. *J Acoust Soc Am*. 2005;117:879–885.