Soul and Musical Theatre.
A comparison of two vocal styles

Hanna Hallqvist[1], Filipa Lã[2], Johan Sundberg[1,3]

[1] University College of Music Education, Stockholm
[2] Center for Social Sciences, Institute of Interdisciplinary Investigation, University of Coimbra
[3] Department of Speech Music Hearing, School of Computer Science and Communication, KTH, Stockholm

Colégio de S. Jerónimo, Largo D. Dinis, Apartado 3087, 3000-995 Coimbra, Portugal
filipa.la@uc.pt

Abstract
Phonatory and resonatory characteristics of non-classical styles of singing have been rarely analyzed in voice research. Here six professional singers volunteered to sing excerpts from two songs pertaining to the Musical Theatre and 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 Theatre, the Soul style was found to be characterized by significantly higher subglottal pressure and significantly greater maximum flow declination rate. The first formant frequency tended to be closer to a spectrum partial in the Musical Theatre style.

Keywords
Voice source, inverse filtering, formant frequencies, non-classical styles

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. 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 non-classical 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 (2008) analyzed aperiodicity aspects of one second long samples taken from recordings of 26 songs that represented six vocal styles. They found statistically significant differences in jitter, shimmer, signal-to-noise ratio and correlation dimension[1]. Björkner (2007), analyzed voice source properties and formant frequencies in male singers performing in the classical operatic and in the musical theatre styles[2] and found significant differences with respect to both voice source and formants. Thus, the opera singers had lower closed quotient, stronger voice source fundamental and lower formant frequencies than the musical theatre singers.

Attempts have also been made to condense the description of different styles of singing. Thurmer (1988) launched the tessiturogram which showed the pitch content of a song in terms of a histogram[3]. Coleman (1987) developed this idea in terms of plots showing SPL as function of percent of the total pitch range of the song[4]. Lykke and associates (2012) applied the method to 206 female conservatory student singers and found that it clearly separated the voices into three basic types, thus suggesting at potential value as a basis for voice classification.

As suggested above, some vocal styles seem associated with specific voice source characteristics, which, in turn, typically are strongly affected by subglottal pressure. Sundberg and Thalén attempted to describe the characteristic combination of this pressure and a phonatory property characteristic in terms of a *phonation map*, showing a measure reflecting phonation type as function of subglottal pressure (henceforth P_{Sub})[5] . It showed clear differences between the analyzed styles, Blues, Pop, Jazz, and Classical.

A phonation map was used also 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)[6]. For instance, Rock showed high P_{Sub}, high F0 and high degree of glottal adduction while Swedish Dance Band showed low values in these parameters, so these styles assumed opposite locations in the phonation map, while 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, non-belt 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 [7]. A similar technique has been found also in some folkloristic styles of singing, like Bulgarian women's singing style and in the Persian singing style called Avaz [8, 9].

The purpose of the present study was to compare two styles of singing in the non-classical repertoire. Soul and Musical Theatre were chosen since these styles appear to be reasonably distinct and can often be professionally performed by the same singers. Our analysis included both voice source and formant frequency characteristics.

Method

Six female singers, age range 22-30 years, volunteered as subjects. They had all studied singing for more than seven years, and had considerable experience of performing both in in the Soul and Musical Theatre genres. They all gave their informed consent to participation.

The singers were asked to sing about 60 s long excerpts from two songs, both written for female singers. One was from the Memphis Soul repertoire *(Son of a preacher man*, lyrics words and music by John Hurley & Ronnie Wilkins) and the other from the Musical Theatre repertoire *(On my own* from *Les Misérables* by Claude-Michel Schönberg), see Figure 1. The total pitch range of both excerpts was similar, G3 – A4 in the former excerpt and A3 – B4 in the latter excerpt, which 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 replacing each syllable in the lyrics with the syllable /pae/.

The subjects were recorded in a sound treated studio (3x4x2,5 m). Using a Soundswell workstation, three tracks were recorded. One track recorded the audio signal, picked up by a head mounted omnidirectional electret microphone (DPA 4065) at a measured distance from the mouth. The microphone signal was amplified by a Symetrix SX 202 Dual Mic Preamp. 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 a Ono Sokki sound level meter LA-210. On a second track was recorded an electroglottograph signal from a Glottal Enterprises MC 2-1 Two Channel Electroglottograph. Oral pressure was picked up 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. The oral pressure during the /p/ occlusion was taken as equal to the subglottal pressure.

Long term average spectra (henceforth LTAS), bandwidth of 400Hz, 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 associated transfer function corresponding to 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 corresponding to the delay of the acoustic signal relative to the EGG. 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, it yields in a separated file fundamental frequency (F0), maximum flow declination rate, defined as the negtive peak amplitude of the derivative of the flow glottogram (MFDR), normalized amplitude quotient, defined as the ratio between the peak-to-peak amplitude of the flow glottogram and the product of MFDR and the period(NAQ), dominance of the fundamental (H1-H2), defined as the level difference between the first and the second partioal 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 Theatre style with significantly higher Leq than that sung in the Soul style, as illustrated in the left panel of Figure 2. The mean difference in Leq amounted to 3.7dB. This seems to support the assumption that the singers used higher P_{Sub} in the Musical Theatre 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 Theatre and in the Soul styles, a non-parametric paired sample test -Wilcoxon- was performed. This particular statistical test was used because data showed a skewed distribution. However, the pressure was significantly higher in Soul [$z = -2.201$; $p = 0.028$], see Figure 3. On average across singers and tones, the pressure in Musical Theatre was 0.84 of the pressure in Soul.

It might seem relevant that the Soul excerpt contained a greater number of short tones than the Musical Theatre excerpt; the average tone duration was 345 ms in Soul and 670 in Musical Theatre, see 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 due to 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 Theatre 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 Theatre style.

With regard to frequencies above 2000Hz, 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 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 Theatre. This supports the assumption that the singers performed the Soul example with firmer glottal adduction, i.e. with a somewhat more pressed type of phonation. However, a statistical test failed to support the same conclusion; a non-parametric paired sample Wilcoxon test, chosen since the data showed a skewed distribution, only revealed significantly higher MFDR values in Soul [$z = -2.201$; $p = 0.028$], see Figure 6.

Figure 7 shows the three lowest formant frequencies used in the inverse filtering analysis for singer 3. In Musical Theatre 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 F1and 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 Theatre, indicating that mostly the formants were further away from a partial in Soul. 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 averages across pitches and singers. In the graph the axes of the ellipses represent ±one 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 Theatre.

Discussion

This study has shown that the six professional singers used higher P_{Sub} and yet produced lower Leq in Soul than in Musical Theatre. This finding supports the assumption that they used stronger glottal adduction in Soul This is in agreement with the trend of the NAQ values to be lower in Soul, 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, i.e., applied more formant tuning, in Musical Theatre. This should increase Leq. Formant tuning has been found to be typically applied in the belt style of singing[7] where loud voice is an important characteristic.

The differences between the phonatory properties of the two styles were rather small both in 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[6]. Rock and Swedish dance Band are phonatorily more extreme than the Musical Theatre 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. The chosen examples were both typical of the repertoire of each style and similar in pitch range, so as to create a musically realistic experimental condition. On the other hand, for each of the pitches analyzed, the musical 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 for 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 also would have caused a great data scatter.

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

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 Theatre. 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 Theatre singing.

Conclusions

Our study has shown that as compared with Musical Theatre, 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 was sung in a somewhat heavier register than Musical Theatre. In addition, the frequency distance between the first formant and its closest partial tended to be narrower in Musical Theatre, which should have contributed to a higher Leq.

References

1. Butte CJ, Zhang Y, Song H, Jiang JJ. Perturbation and nonlinear dynamic analysis of different singing styles. *Journal of Voice.* 2008;23:647–652.
2. Björkner E. Musical theater and opera singing—Why so different? A study of subglottal pressure, voice source, and formant frequency characteristics. *Journal of Voice.* 2007;22:533–540.
3. Thurmer S. The tessiturogram. *Journal of Voice.* 1988;2:327–329.
4. Coleman RE. Performance demands and the performer's vocal capabilities. *Journal of Voice.* 1987;1: 209-216.
5. Thalén M, Sundberg J. Describing different styles of singing: A comparison of a female singer's voice source in "Classical", "Pop", "Jazz" and "Blues". . *Logopedics Phoniatrics Vocology.* 2001;26:82-93.
6. Zangger Borch D, Sundberg J. Some phonatory and resonatory characteristics of the Rock, Pop, Soul, and Swedish Dance Band styles of singing. *Journal of Voice* 2011;25:532–537.
7. Sundberg J, Thalén M. Respiratory and acoustical differences between belt and neutral style of singing. *Journal of Voice.* 2015;29:418–425.

8. Henrich N, Kiekac M, Smith J, Wolfe J. Resonance strategies used in Bulgarian women's singing style: A pilot study. *Logopedics Phoniatrics Vocology* 2007;32:171-177.

9. Biglari HM, Sundberg JJASA, (). Timbral and melodic characteristics of Persian and Kurdish singing *J Acoust Soc Am* 2012;131:3377.

Discrimination of Three Basic Female Voice Types in Female Singing Students by Voice Range Profile-Derived Parameters H. Lycke a W. Decoster a A. Ivanova b M.M. Van Hulle c F.I.C.R.S. de Jong Folia Phoniatr Logop 2012;64:80–86

Figure captions

Figure 1. Music excerpts used in the experiment representing Musical Theatre and Soul (left and right scores, respectively. The tones analyzed are marked with circles.

Figure 2. Left panel: Equivalent sound level measured in the Soul and in the Musical Theatre excerpt. Numbers refer to singers.
Right panel: Subglottal pressure used for identical pitches when occurring in the Soul and in the Musical Theatre excerpt.

Figure 3. Box plot of the subglottal pressure (P_{Sub}) used in the Musical Theatre and Soul excerpts.

Figure 4. Long-term-average spectra of the Musical Theatre and the Soul excerpts (gray and black curves, respectively).

Figure 5 Values of pulse amplitude, MFDR, Q_{Closed}, H1-H2 and NAQ, used by the singers for identical pitches in the two excerpts.

Figure 6 Box plot of the MFDR used in the Musical Theatre and Soul excerpts

Figure 7. F1, F2 and F3 of the vowel /ae/ observed when Singer 3 sang the Musical Theatre and the Soul example. The thin dotted curves show the frequencies of the spectrum partials.

Figure 8. Left panel: Average normalized distance between the F1 and its closest partial in the indicated singers' performances of the Musical Theatre and Soul examples. Gray and black columns refer to the Musical Theatre 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 theatre and Soul (diamond and circle). The gray and black ellipses reflect the standard deviations for the same styles, respectively.