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. 31, No. 2, pp. 229–235
0892-1997

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

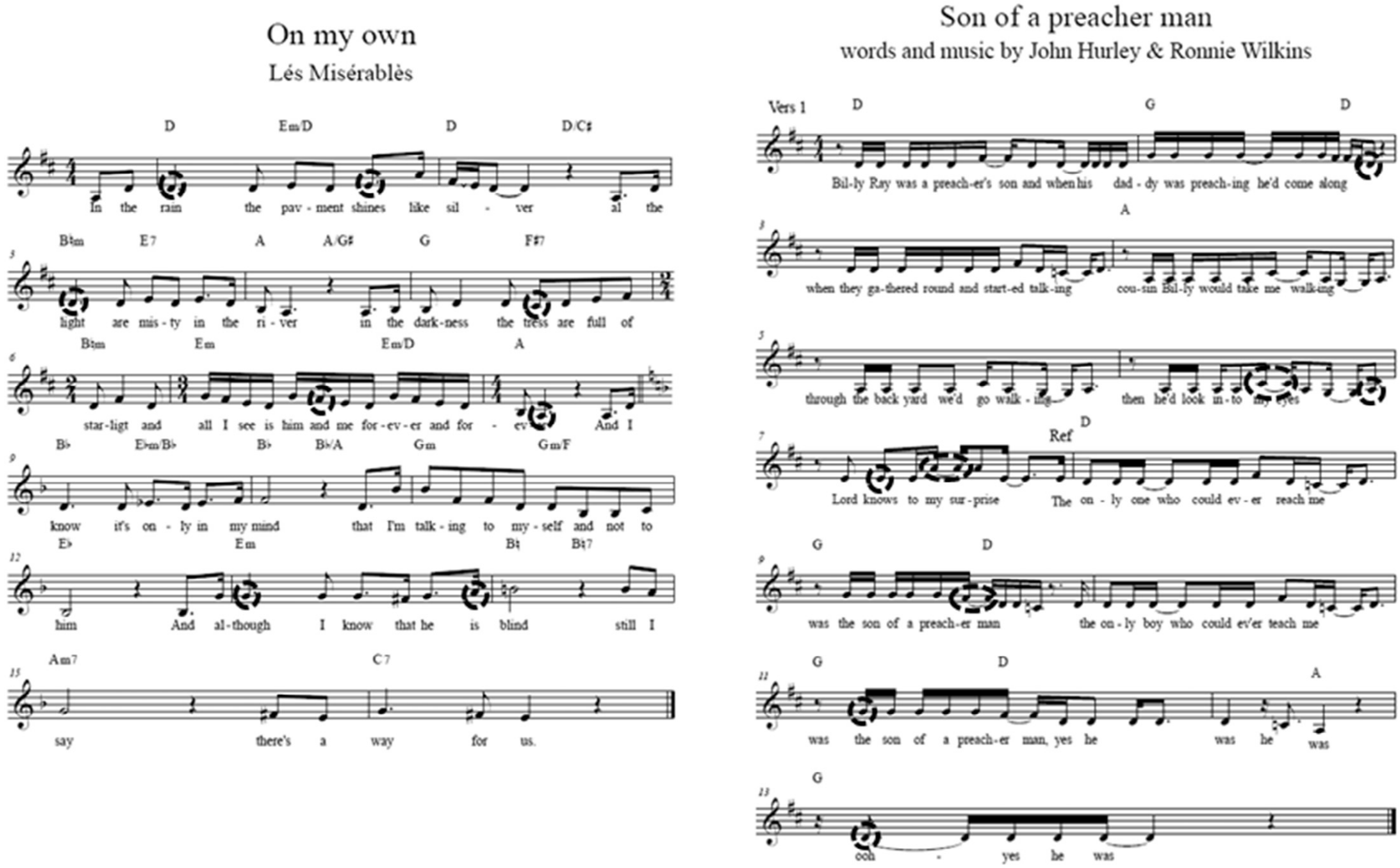

FIGURE 1. Music excerpts used in the experiment representing musical theater and soul (left and right scores, respectively). The tones analyzed are marked with *circles*.

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.