Vocal Tract Shaping and Dark Tone Quality
An Investigation Using Real-Time MRI

Elisabeth R. Lynn*, Shrikanth S. Narayanan, and Adam C. Lammert***
***Worcester Polytechnic Institute, **University of Southern California**

Introduction

Vocal pedagogy has identified the tone quality "dark" as an aesthetically ideal quality of Western classical voice performance. Previous studies indicate that darkness is associated with lowered formant frequencies[5], perhaps stemming from lowered larynx and widened pharynx[1]. This study uses rtMRI data to examine the production of different tone qualities, with the objective of confirming previous findings, and identifying additional articulatory correlates of dark tone quality.

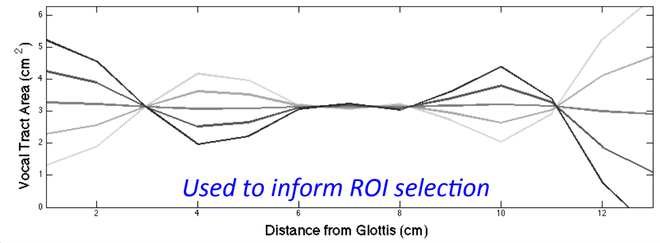
Subjects & Data Acquisition

Experiment: Four sopranos trained in the Western Classical tradition sang portions of "Susanna's aria" from Mozart's Marriage of Figaro. Subjects were instructed to sing in three separate tone qualities: normal performance style (instructions identified as "ringing"), especially dark (identified as "covered"), and especially non-dark (identified as "bright").

rtMRI: Data were collected on a GE Signa 1.5T scanner, using the protocol described by Narayanan[3]. Midsagittal images were reconstructed at an effective rate of 22.41 fps, with a spatial resolution of 2.9 × 2.9 mm/pix. Sychronous audio was recorded at 20 kHz, and denoised using the method described by Bresch[2].

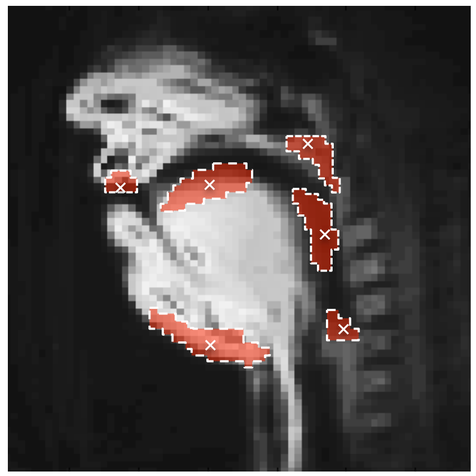
Method: Artic/Acoust Simulations

Formant frequencies were computed for randomly-generated vocal tract shapes. "Dark"-type formant lowering was correlated with vocal tract shape using Canonical Correlation Analysis.



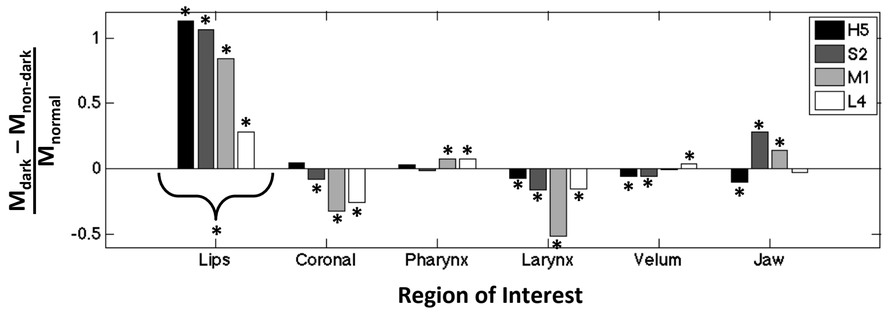
Method: Region of Interest

The mean image for each subject and tone quality was created, and used to select and form ROIs based on the method described by Lammert[1]. ROIs were used to examine vocal tract kinematics



Subject M1 as represented by the mean of all images acquired during her dark aria performance. The six ROIs are superimposed in red, with the selected pixels used to form each region shown as a white 'x'. ROIs correspond to labial, coronal, pharyngeal regions, a region in the larynx immediately superior to the glottis, as well as regions capturing velum raising and jaw lowering.

Results: Articulatory Analysis



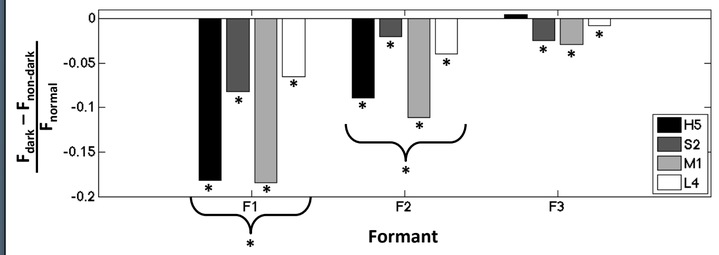
Differences in vocal tract shaping associated with dark tone quality
Within-subjects: Differences observed most regions
Across-subjects:

- Labial narrowing, statistically significant difference
- Larynx widening, substantial and consistent, but not significant#
- Coronal widening, substantial and consistent, but not significant#

#*These patterns are also predicted by the physical acoustic simulations, and may therefore represent a real effect that would achieve statistical significance given a larger number of study participants.*

Method: Acoustic Analysis

Formant frequencies 1-3 were extracted from voiced sounds, as determined by pitch analysis, using Praat. The expected lowering of formants was observed.



Used to inform physical acoustic simulations

Conclusions/Future Work

Conclusion #1: Results differ in from traditional accounts, which emphasize pharyngeal widening

Conclusion #2: Pharynx does seem to play a role, but that role is highly variable across individuals, perhaps reflecting individualized training, or morphology

Conclusion #3: Labial aperture may be the primary driver of dark tone quality, perhaps opposed to "twang" which has been related to labial widening[4].

Future work: Additional ROIs will be analyzed and additional details of tone quality production kinematics. Results may be used to consider the impact that tone quality production has on vocal health.

References

[1]Bloothooft G & Plomp R (1986) Spectral analysis of sung vowels. III. Characteristics of singers and modes of singing. JASA, 79(3), 852-864.

[2]Bresch E, Nielsen J, Nayak K & Narayanan S (2006) Synchronized and noise-robust audio recordings during realtime magnetic resonance imaging scans. JASA, 120(4), 1791-1794.

[3]Narayanan S, Nayak K, Lee S, Sethy A & Byrd D (2004) An approach to real-time magnetic resonance imaging for speech production. JASA, 1771-1776.

[4]Story BH, Titze I R & Hoffman EA (2001) The relationship of vocal tract shape to three voice qualities. JASA, 109(4), 1651-1667.

[5]Sundberg, J (1970) Formant structure and articulation of spoken and sung vowels. Folia Phoniatrica et Logopaedica, 22(1), 28-48.