Acoustical Correlates for Seven Styles of Singing

Laura Anne Bateman
University of Victoria, McPherson Library, Victoria, BC V8W 3H5

Legend has it that one day Garcia was strolling down the street. He was thinking of his sister, a famous opera diva and wondering how people could create such beautiful sounds. Could he find a way to look inside the instrument and see how it worked? Just at that moment, the sun caught the tip of his walking stick and the light was reflected onto his face. The idea of the laryngeal mirror was conceived and voice science born. These early efforts in voice research took place in the first half of the 19C. By the late 1800's incandescent light was used with stroboscopy to provide more answers, but it wasn't until the development of microtechnology such as fiberoptics, digital analysis and imaging techniques in the 20C that voice science has made the greatest strides.

This century William Vennard, was one of the American pioneers in singing research. He looked at the acoustics and physiology of the classical singing voice. (Vennard, 1967) Johan Sundberg, has also provided much information about the classical singing voice from his continuing research at KTH Voice Research Centre in Sweden. (Sundberg, 1987). The first singing qualities studied were "trained" singers vs "untrained" singers. These early studies provided some answers but, as Jo Estill, an American voice teacher and researcher pointed out in a recent email correspondence "...they are trained or untrained. But trained to do what." (Estill, 1997). Sundberg has since become more specific with the type of voice quality and expanded his focus to include choral, belt and twang qualities. Jo Estill has broadened the scope further to include six basic voice/singing qualities: speech, falsetto, sob, opera, twang, and belt. (Estill, 1996) Ingo Titze has also written extensively on the physics of the singing voice. (Titze, 1994) However, many of the voice qualities people use in contemporary popular music remain somewhat of a mystery acoustically and physiologically.

Voice quality research also has another branch in the field of Linguistics. The most comprehensive taxonomy comes from John Laver, University of Edinburgh. (Laver, 1980). Some is quite specific to the speaking voice, but the concepts can be adapted to the singing voice. Voice quality in speech "is the auditory colouring of an individual speakers voice" (Laver, 1980 pg 1) and involves voice quality settings. For example, a speaker might have a tendency to keep the lips rounded throughout speech. This colours the speaker's utterance. Laver further divides voice quality settings into laryngeal and supralaryngeal settings. These are present for the duration of the speech sequences. For example, the previous subject may have used modal voice with labial protrusion. Modal voice would refer to the laryngeal setting and labial protrusion would refer to the supralaryngeal setting. These settings have corresponding diacritics. In singing, the same concepts can also be applied. For example, using Laver's taxonomy, a belter might use modal voice with a pharyngealized tongue body notated as a capital V with two horizontal wavy lines through it. This describes a physiological configuration, but it is also an auditory label with an acoustic correlate. Phonetic taxonomy also describes speech as having specific units called segments (vowels and consonants). These segments have articulatory, acoustic and perceptual features. As with a speech dialect, segmental features can be analysed in singing style for acoustic correlates. For example, the pop style may use a less aspirate "t" than the classical style. The classical "ah" may be darker than the bright country "ah". This presentation will focus on the acoustic correlates of voice quality settings however, segmental features may also be discussed.

Jo Estill refers the combination of features as a "recipe" and trains singers to make a series of physiological adjustments to obtained a prescribed voice quality. Her list of parameters include: onset, constriction and retraction of the false vocal folds, vocal fold plane, vocal fold mass, laryngeal tilt, soft palate control, anchoring, pharyngeal width, pharyngeal length, tongue control aryepiglottal constriction. Dang and Honda (1996) also clarified the strong effect that the piriform sinuses have in voice quality. Physiological correlates for this presentation are currently under investigation and will not be discussed, however, acoustic results may be compared to the existing literature.

Much acoustic description of the voice still uses a model developed by Fant. (Fant, 1960) He described the voice using a linear model involving a source and three filters. He assumed the glottal source to be white noise filtered at the glottis to produce a source function which decreases by 12dB per octave in strength, the vocal tract as a single tube closed at one end which acts as a second filter producing the odd numbered formants at approximately 500Hz, 1500Hz, 2500Hz etc. and the radiation effect as a third filter with a function increasing by 6dB per octave. This is the model which analysis algorithms such as LPC use so it is worth mentioning. Whether this model is "good enough" is always a contentious issue and the development of other models (Titze, 1997) and analysis algorithms is ongoing. The analysis methods used in this investigation are Spectrograms generated using Fast Fourier Transform algorithm in 3D graphic format (CSL software), and LPC using the auto-regressive model and autocorrelation (solved using the Levinson function). Transfer function plotted using z-transform digital filter frequency response. (Matlab software). Decisions as to sampling rate, order number, windowing, preemphasis and other settings will be discussed.

The data for this presentation comes from a case study of Lisa Popeil, a singing teacher and performer, who can produce seven voice quality settings: pop, jazz, R&B, belt, opera, legit, country. Ms. Popeil has been a Los Angeles studio musician for many years and her singing styles have been heard on radio, advertising, cd, video and in live performance with such notable artists as Frank Zappa and Weird Al Yankovic; therefore, the samples obtained are believed to approximate the norm for the singing styles under consideration. This presentation will focus on formant placement. Steady state and running sample vowels were sampled and analysed to determine formant frequencies and approximate energy in each formant. The values were then compared. Similarities and differences between the styles were noted. The styles will be related to labelling schemes by Jo Estill and John Laver. Perceptual testing has not been included in this investigation.

This type of research has been a great help to singing teachers as well as engineers for producing voice enhancement and education-

al tools for singers.

REFERENCES:

DANG, J. and HONDA, K. (1996). Acoustic characteristics of the piriform fossa in models and humans. J.Acoust.Soc.Am. 100, 3374-3383.

ESTILL, J., FUJIMURA O., SAWADA A. AND BEECHLER K. (1996), Temporal Perturbation and Voice Qualities in P. Davis and N Fletcher (Eds) Vocal Fold Physiology: Controlling Complexity and Chaos, V.9 Ch. 16 237-252 San Diego: Singular Publishing Group.

ESTILL, J. (1997) Email correspondence Feb. 5/97.

FANT, G. (1960), Acoustic Theory of Speech Production. The Hagues:Mounton.

LAVER, J (1980). The Phonetic Description of Voice Quality. Cambridge: Cambridge University Press.

TITZE, I. (1994). Principles of Voice Production. New Jersey: Prentice Hall.

TITZE, I. and STORY, B. (1997) Acoustic Interactions of the Voice Source with the Lower Vocal Tract. J.Acoust.Soc.Am. 101(4), 2234-2243.

SUNDBERG, J. (1987). The Science of the Singing Voice.. Illinois: Northern Illinois University Press.

VENNARD,W. (1967). Singing The Mechanism and the Technic.. New York: Fischer.