ISCA Archive
http://www.isca-speech.org/archive

First International Conference on
Spoken Language Processing (ICSLP 90)
Kobe, Japan
November 18-22, 1990

A STUDY ON RESPIRATORY AND GLOTTAL CONTROLS IN SIX WESTERN SINGING QUALITIES: AIRFLOW AND INTENSITY MEASUREMENT OF PROFESSIONAL SINGING

Jo Estill, *Noriko Kobayashi, **Kiyoshi Honda, and **Yuki Kakita

Shiki Theatrical Company, *ATR Auditory & Visual Perception Research Labs.,
and **Kanazawa Institute of Technology

ABSTRACT

The purpose of this study is twofold: (1) to examine respiratory controls when voice quality was specified, (2) to rate the relative efficiency of each of the voice qualities recorded. One female professional singer, who can consistently control six qualities was the subject of the experiment. Multiple repetitions were made of each quality at four fundamental frequencies: 192, 294, 392, and 587 Hz, as airflow, sound pressure levels, and electroglottographic measurements were recorded. In most of the qualities, air flow rate and acoustic intensity showed high correlation. Glottal efficiency, computed from these aerodynamic and acoustic measures, divided the six singing qualities into two groups: four loud qualities and two soft qualitities. In the loud group, glottal efficiency increased with higher frequencies. The soft qualities, on the other hand, showed maximum efficiency in the middle frequency range.

INTRODUCTION

In many past studies of laryngeal function, subjects have phonated in a "normal" speech quality, sometimes called "modal" speech, implying there is only one voice quality in normal speech. However, we hear many voice qualities in the "normal" speech around us, and there is little variation required or expected from any individual speaker. When intensity variations were measured in these studies, the subjects generally increased the effort in whatever qualities they habitually used. Regardless of the individual variations of voice quality, a relatively limited range of physiological capability is required for the production of speaking voice.

The ultimate goal of our research is to enlarge our understanding of the physiological potentials of human laryngeal function for voice production. This goal can be achieved by studying a variety of voice modes beyond those normally used by the average speaker. In order to make accurate comparisons of the physiology among these qualities, the qualities should be produced by the same vocal instrument. For this purpose, singers and actors, trained to make different qualities, make the best subjects to provide accurate comparison of physiological differences in voice production. Therefore, we chose to examine one such subject and six voice qualities; Speech, Falsetto, Sob, Twang, Opera, and Belting.

The objective in the study was to observe the capabilities of a single voice to vary intensity, not by increasing the effort in one quality, but by changing qualities remaining always at a most confortable effort level in each voice modality. The purpose here was to observe airflow and laryngeal function when the single subject controlled the laryngeal and vocal tract configuration appropriate for each quality (Colton and Estill, 1981; Honda and Estill, 1986).

METHOD

One of the authors, a female professional singer and voice trainer, served as the subject of the experiment. The subject produced six western singing qualities: Speech, Falsetto, Sob, Twang, Opera, and Belting, as described in Table.1. These qualities are defined not only by perceptual impression but also by somatosensory characteristics. Each quality was produced at four different frequencies, i.e., 196, 294, 392, and 587 Hz, in a sustained vowel /i/ with the voice quality consistent for all frequencies, while effort was concentrated on controlling the configuration, position and muscle tension of the various parts in the the laryngeal and supra-laryngeal organs. Ten repetitions were made of each quality and frequency.

Air flow rate and sound pressure level were measured using the PS-77 phonatory function analyzer (Nagashima Medical Instruments Co., LTD.), which is designed for the clinical evaluation of voice disorders. A silastic ventilator mask was connected to the sensor assembly inlet of the PS-77. The subject phonated with the mask held on the face to cover the mouth and the nostrils . Electroglottographic signals were simultaneously recorded using the Laryngograph (Laryngograph, LTD.). These signals were recorded on a digital data recorder. Ten repetitions of each quality at each frequency were sampled for data analysis.

The data analysis was performed for the stable portions of those sampled data. The average levels were manually extracted to measure mean airflow rate U[ml/sec] and sound pressure level I[dB]. Values for mean airflow rate were converted into logarithmic values to be compared in the same decibel scale as sound pressure level. An equiva-

TABLE 1
Voice quality, musical genre, and speculative physiological
description of voice modes recorded

Speech	Popular singing, Recitatives	Totally relaxed vocal tract, each tone initiated with a slight glottal attack.
Falsetto	English choir boy	Totally neutral and relaxed vocal tract, initiated by blowing the tone "on the breath." The vocal folds are presumably thin and the cover is stiff due to passive tension.
Sob	Classic/concert	Low larynx, expanded vocal tract, thin folds, body "anchored," and a sensation and sound of a repressed sob.*
Twang	Country-Western	High larynx, tightened aryepiglottic sphincter, open velopharyngeal port, little effort at the vocal folds.
Opera	(with "squillo" or "singing formant")	Tightened aryepiglottic sphincter, somewhat lowered larynx, "anchored."
Belting	Broadway Theatre	High energy, high intensity yelling, high tongue, high larynx, "anchored."

*"Anchoring" is a term used to describe a contraction of the muscles in the head, neck, or the use of larger extrinsic muscles in the interest of finer control over intrinsic muscles. From preliminary research and an unpublished paper, "Zeroradiography and 'anchoring.'" in 1989 by Estill J., and Bagnall, A. (Children's Hospital, Adelaide, South Australia).

lent glottal efficiency was then calculated from these two kinds of values for all the data.

Efficiency E=I-U,
where, U=20log(U/U0) and U0=1 [ml/sec] (Kakita, 1987)

RESULTS

In Fig.1 is displays all the data showing relationships between airflow rate and sound pressure level [SPL], both plotted in logarythmic scales. In each diagram is shown the data distribution of the two parameters for the six qualities at one selected frequency.

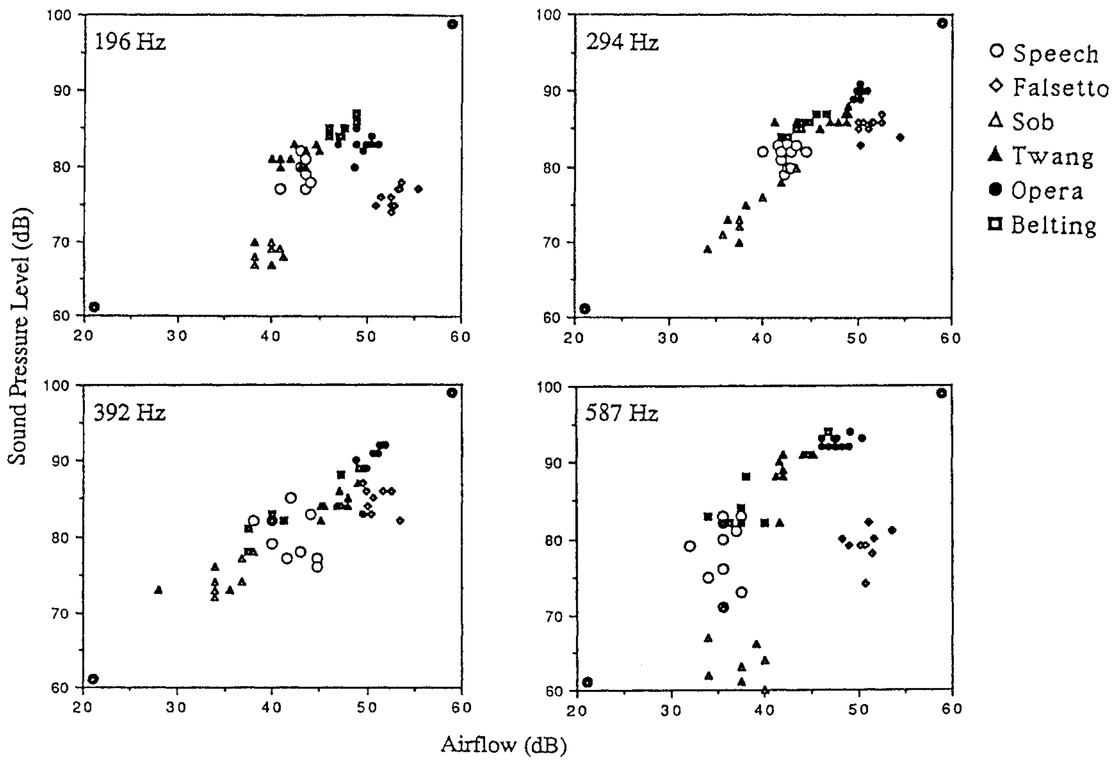

Fig.1. Scattered diagrams to show the relationships between airflow rate and sound pressure level. Each diagram contains the data for ten repetitions in all qualities at one frequency. Logarithmic conversion was performed for the airflow rate to be shown in a dB scale.

In all the data, SPL tended to increase as airflow rate increased. A linear relationship was observed between the two parameters, particularly at the two middle frequencies, 294 Hz and 392 Hz. In the group of loud qualities, Twang, Opera, and Belting, this relationship was preserved at all frequencies. In these louder qualities, airflow rate and SPL tended to be higher than speech in various degrees. Among them, Opera had the highest air flow rate at all frequencies. Other loud qualities showed variations. Twang quality overlapped with Speech at the lowest frequency 196 Hz. As the frequency increased, however, the airflow rate tended to approach that of Opera. Belting had a higher air flow rate at 196 Hz than Twang, but at the highest frequency, 587 Hz, Belting airflow overlapped with Speech.

In the two soft qualities, Falsetto and Sob, the data distributions were different from other qualities. In both Falsetto and Sob qualities, the SPL was lower than that expected from the airflow rates observed in louder qualities. However, there was a distinct difference between Falsetto and Sob in airflow rate and SPL. Sob was consistently lower than Falsetto in both dimensions at all frequencies.

Changes in SPL, airflow rate, and glottal efficiency are shown in Fig. 2(a), (b), and (c) respectively, for six qualities at four different frequencies. Averaged values obtained from ten repetitions of six qualities are used in each single plot. The values for average SPL, as shown in Fig. 2(a), stayed at relatively constant level within each quality. At the highest frequency of 587 Hz, SPL levels tended to separate to show the characteristics of those qualities. Opera had the highest SPL of 92.5 dB, and Sob, the lowest, of 63 dB. The dynamic range observed in the present data was approximately 30 dB.

Average airflow rate, shown in Fig. 2(b), tended to decrease gradually with frequency in most of the qualities. These changes occurred in parallel to each other within a range of approximately 15 dB.

In the glottal efficiency, shown in Fig. 2(c), differences were observed between loud and soft quality groups. Loud qualities have higher glottal efficiency than soft qualities. In loud qualities, Twang, Opera, and Belting, as well as in Speech, glottal efficiency tended to increase as frequency increased. However, the data did not show a linear relation with frequency, because in most of the louder qualities, glottal efficiency at 392 Hz was slightly lower than that at 294 Hz. In soft qualities, glottal efficiency increased with frequency up to 392 Hz, but at 587 Hz, the efficiency in both soft qualities decreased. Variations in glottal efficiency were larger at the lowest and the highest frequencies. The dymanic ranges at the lowest frequency were about 15 dB, while at the highest frequency the range was about 20 dB. The most efficient quality was Belting at 587 Hz (47 dB), while the least efficient quality was Falsetto at 196 Hz (22 dB). All qualities fell within an efficiency range of 25 dB.

The EGG signals for Speech, Falsetto, and Belting at all the frequencies are shown in Fig. 3. The data for Falsetto and Belting are chosen to represent the soft and loud groups respectively with a comparison of that for Speech quality.

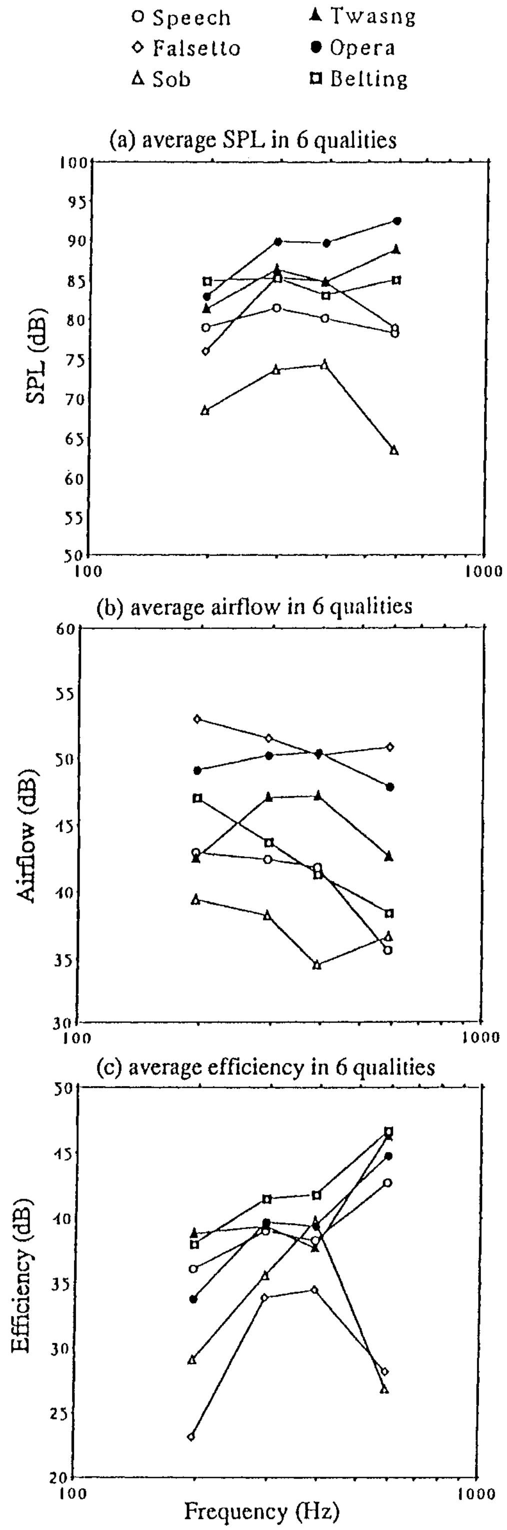

Fig. 2. Average values for (a)sound pressure level, (b)airflow rate, and (c)glottal efficiency, in all qualities at four frequencies. Each data point represents the average value for ten repetitions.

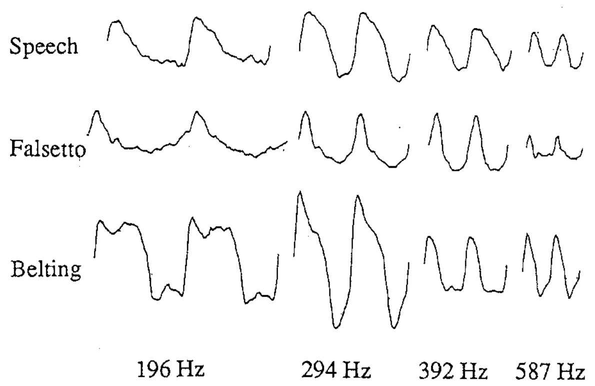

Fig. 3. Examples of EGG waveforms showing typical patterns of the changes of the vocal fold contact area[VFCA], in Speech (top), Falsetto (middle), and Belting (bottom). Increase of the VFCA is shown as the upward movement of the signal.

Two cycles of typical waveforms were extracted from a stable portion of the signals in one repetition. By viewing the changes in the vocal fold contact area [VFCA] at each frequency, we noticed that the most remarkable difference was observed across qualities at 196 Hz. Falsetto has a large open quatient [OQ], while Belting has a smaller OQ. A similar observation may be possible for the data at the other frequencies. However, the data for a middle frequency, 394 Hz, had less significant difference for temporal characteristics.

DISCUSSION

At all frequencies there was a range of data among these six qualities, suggesting that whatever parameter is being measured, airflow, sound pressure level, or efficiency, that the result will vary depending on the voice quality used by the subject.

The most striking differences in the soft qualities were those between Falsetto and Sob, where Falsetto had the highest airflow, Sob the lowest. This finding may be of interest to speech pathologists who must retrain patients with functional voice disorders. If the major goal of voice therapy for such patients is to find the least traumatic use of the voice, it would appear that Sob with the least airflow and VFCA would be of more benefit than Falsetto quality with the highest airflow and some VFCA. The difference in production between the two qualities and the reason for the reduction in airflow with Sob quality may be attributed to the feature, "anchoring," where the effort was made to control the airflow.

If the intent of the speaking mode is a voice that has a higher efficiency, then Twang and Belting rate slightly higher than Speech with Twang the more efficient.

Of interest to singers is the data showing that for the most efficient qualities, Belting, Twang, Opera, and Speech, that for the highest frequency, airflow decreased. As airflow decreased for Belting, Twang, and Opera, the SPL and efficiency increased.

The EGG data for loud and soft qualities coincide with the result for glottal efficiency. A quality having a larger OQ demonstrates a lower glottal efficiency. From the result, we may note easily that intensity variation results from glottal and respiratory controls that alter the source function. Also, we can presume quite naturally that such controls are coordinated with the activity of supra-laryngeal articulators for adjusting the vocal tract shape to enhance intensity variations.

Finally, in these data one will note that all the qualities could be produced at both ends of the F0 range. There was no division into "registers" at the perceptual level. At the mechanical level, however, one might postulate a difference at 294, 392, and 587 Hz, suggesting there might have been a change in mechanical action to make the higher frequencies.

SUMMARY

One subject was recorded producing six qualities testing differences in respiratory control (airflow), laryngeal vocal fold contact area (EGG), and intensity (SPL) at four frequencies, 196, 294,392, nand 587 Hz.

Respiratory differences were most evident in the two soft qualities, Falsetto and Sob, with Falsetto exhibiting the highest airflow and Sob, the least, at all frequencies. In general, airflow tended to decrease at the highest frequency in all qualities.

In the sound pressure level and efficiency of the six qualities tested, Belting, Twang, and Opera followed by speech were the most efficient at both the lowest and highest frequencies. Except for Sob at 392 Hz, the two soft qualities, Falsetto and Sob were least efficient at all frequencies.

REFERENCES

Colton, R.H. and Estill, J. (1981). Elements of voice quality: Perceptual, acoustic and physiological aspects. Chapter in N.A. Lass (Ed.) *Speech and Language: Advances in Basic Research and Practice*, New York: Aademic Press, pp311-402.

Honda, K. and Estill, J. (1986). Pitch control in six voice qualities: an EMG study. *Proceedings of the XXth Congress of the International Association of Logopedics and Phoniatrics*, Tokyo, pp 326-327.

Kakita, Y. (1987). Measures and displays representing phonatory ability. In T. Baer, C. Sasaki, and K. S. Harris (Eds), *Laryngeal Function in Phonation and Respiration*, Boston: Little Brown and Company, pp 448-462.