ON THE EFFECTS OF VOCAL TRAINING ON THE SPEAKING VOICE QUALITY OF MALE STUDENT ACTORS

Timo Leino & Päivi Kärkkäinen,
Institute of Speech Communication and Voice Research, University of Tampere, Finland.

ABSTRACT

Male student actors were given as an extra element in their ordinary voice training a special training period of eight months for strengthening of the overtones especially around 3.5 kHz. A spectrum analyzer was used troughout for visual feedback. The reading samples after training had a less steep slope in the long-term average spectrum and the peak at 3.5 kHz was in some cases more prominent. Samples were rated to sound better. After two years the changes still existed.

INTRODUCTION

In earlier studies Leino [1, 2] found that in the long-term average spectra (LTAS) from text reading samples good and poor male actors' voice qualities differed from each other the former having a less steep slope and a strong peak at 3.5 kHz. This peak was named an "actor's formant" by the author. Strengthening of the peak and all overtones was chosen for a goal of a special voice training period included in the ordinary voice training of male student actors. The present article summarizes the results of this project.

MATERIALS AND METHODS

After one year of ordinary voice training seven male student actors on a four year training course for professional actors received an extra 8 month voice training period. A training session of 20 - minutes was given once a week. The ordinary voice training continued at the same time. Special attention was payed to strengthening of the overtones especially around 3.5 kHz. Therefore a real time spectrum analyzer (Spectral Dynamics SD301/SD309) was used to give visual feedback throughout the training sessions. The vocal exercises consisted of nasal-vowel syllable strings produced aiming at a clear, bright, well projecting voice quality.

Before and after the training period the same prose extract of about one minute was recorded in a sound-treated studio using Revox A700 tape recorder and Electrovoice RE11 microphone 40 cm from the subject's mouth. The loudness was kept at normal reading strength on both occasions. Long-term average spectra were made from the text reading samples with a Hewlett-Packard signal analyzer (3561A). A four-hundred point narrow band FFT analysis was used. The frequency span was 10 kHz. Voiceless segments were excluded. The time record length was 40 ms. The display resolution was 25 Hz. The Hanning weighting window with the frequency band of 37.5 Hz was used.LTAS were made of individual samples and averages calculated from individual LTAS with a microcomputer. LTAS were compared according to the slope. For this purpose the spectra were studied on a relative scale where the strongest amplitude peak was given the value zero.

The samples recorded before and after training were played in random order to various groups of listeners including university students and theater and speech professionals.Text reading samples by the students were re-recorded and analyzed after two years of ordinary voice training, where no special attention was paid to strengthening of the overtones and the spectrum analyzer was no longer used in training.

RESULTS AND DISCUSSION

Figure 1 compares average LTAS (a) before and after special voice training and (b) before the special training period and two years after it was over and only ordinary voice training was given. The figures show only the frequency range 0-4 kHz, since no significant changes were observed above this. It can be seen that on average the spectral slope became less steep and the peak around 3.5 kHz became somewhat more prominent after the special training period (Fig. 1 a). After two years these characteristics were mainly still seen but weaker (Fig. 1 b).

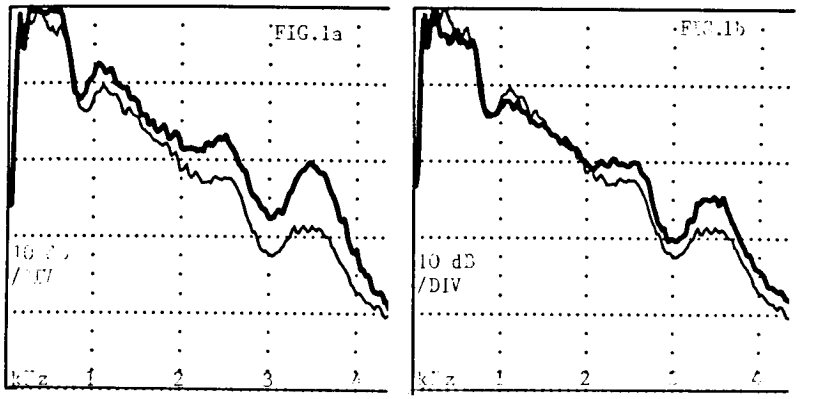
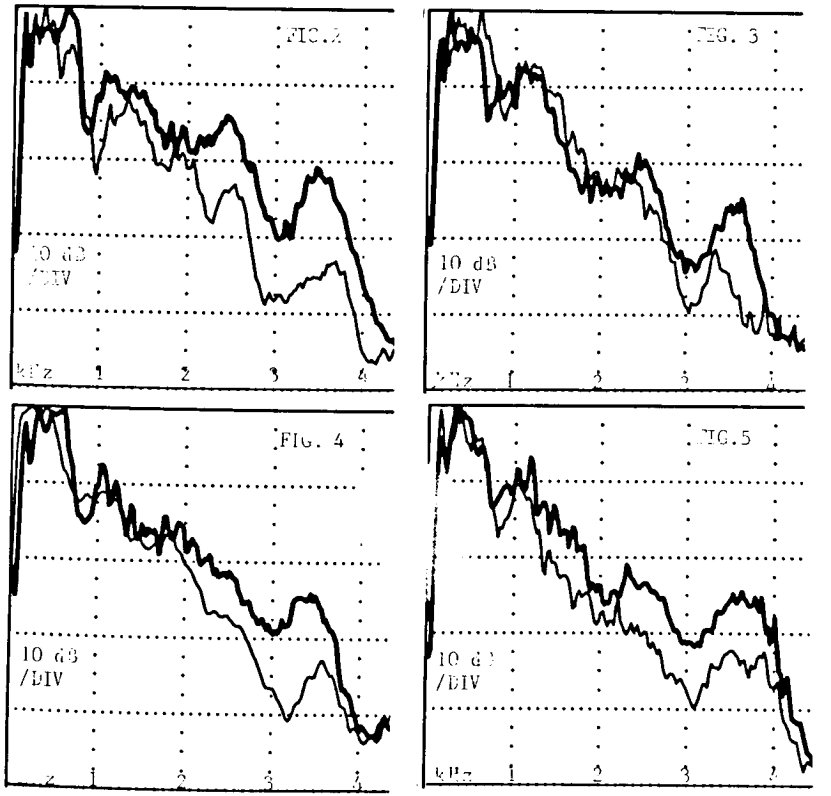


Figure 1. Average LTAS (a) before (thin line) and after (thick line) the eight-month special training, (b) before (thin line) and two years after the special training (thick line).



Figures 2-5. Changes in the LTAS of four individual student actors after the eight-month special voice training. Thin line = before, thick line = after.

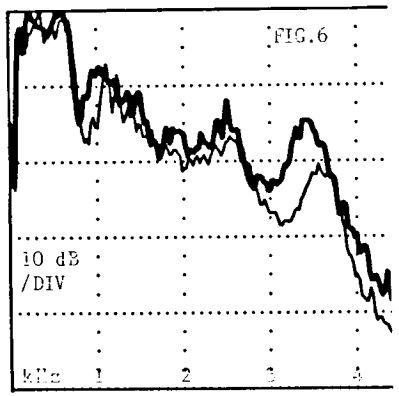
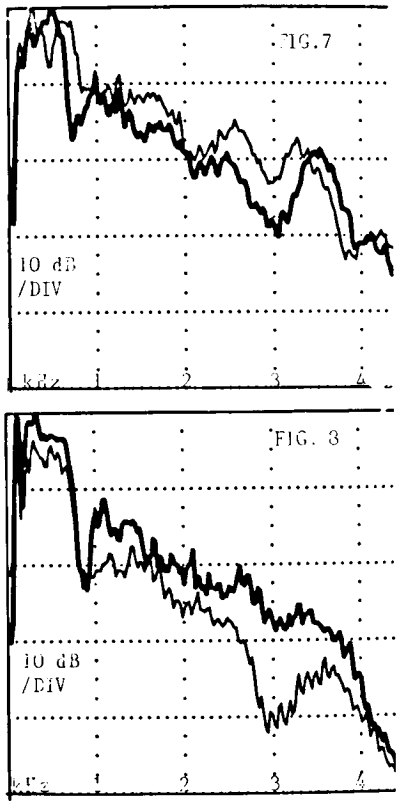


Figure 6. LTAS of one student, whose voice quality after the special voice training (thick line) received conflicting evaluations.

Figures 2-8 compare LTAS before and after the training period for each student. After the training in all but one case the slope became less steep and the peak around 3.5 kHz slightly more prominent in four cases. Only in two cases the peak of 3.5 kHz was as strong as the peak of 2.5 kHz or stonger than that, which according to the earlier findings of Leino [1]. seems to be one characteristic feature of a very good voice.

The text samples read after the training period were in the listening tests evaluated to sound better. Only in one case (Fig. 6) the listeners disagreed. Obviously that student, who had before the training period already had strong overtones and a prominent peak at 3.5 kHz, had tried too hard to make his voice even better with the result that the overtones became too strong for some of the listeners. In his LTAS the difference between the strongest spectral peak and the peak at 3.5 kHz was only 14 dB while for other students this difference was 20-25 dB after training.

Leino [2] has earlier found that if a speaking voice sample is altered by filtration, the voice quality is evaluated to be better if the amplitude difference between the 3.5 kHz peak and the strongest spectral peak is about 15-30 dB. The voice quality rating is impaired both when this difference increases and when it decreases.



Figures 7 and 8. LTAS of two students before (thin line) and after (thick line) the training period.

Another student (Fig. 7) behaved in the opposite way. He also had already rather strong overtones, but he did not try too hard to make them stronger. Instead, most likely he tried to change his voice quality in the exercises through resonatory changes by only increasing the prominence of the 3.5 kHz peak. Similarly the student whose LTAS can be seen in Fig. 8 seems to have changed his resonatory setting, which has contrastively led to the disappearance of the peak.

In general the spectral changes related to voice training and improvement of voice quality may be explained from the basis of both phonatory and resonatory changes. The spectral slope is known to be related to the glottal closing speed so that the increasing closing speed gives a less steep slope [3]. Also Frøkjær-Jensen & Prytz [4] and Wedin et al. [5] have found that vocal training decreases the spectral tilt and increases the perceptual "sonority" of the voice quality. However, the clear peak around 3.5 kHz which has been found to be one characteristic of a very good male voice quality [1-2], seems also to require resonatory bases. The valleys separating the peak from its surroundings suggest that it is formed by a formant or a cluster of two or more formants, most likely F4 and F5. This frequency range has been regarded as more prone to resonatory changes than phonatory ones [6]. Nolan has also found a clear peak at 3.5 kHz in his voice [7]. This peak was especially prominent in creak and creaky voice, largely absent in falsetto, in whispery voice as well as in raised or lowered larynx voice and totally absent in whisper. Nolan considers the possibility that this peak is a phenomenon similar to singer's formant, which, according to Sundberg, [8] is a result of laryngeal resonance arising when the cross-sectional area of the outlet of the larynx tube is sufficiently different from the cross-sectional area of the pharynx.

In conclusion, the results suggest that by vocal exercising it is possible along with consciously set goals to strengthen the overtones and also in some cases increase the prominence of the peak around 3.5 kHz, and that this change in voice quality is perceptually evaluated as positive. Spectrum analyzer seems to be a useful aid in visualising the aims of the exercises to the students. This may make the learning process faster and increase the motivation of the students. The results also show the limits beyond which the strengthening of the spectral peaks and overtones in general is no more perceptually acceptable and most likely also from the voice hygienic point of view questionable. The importance of individually set goals for every student must be emphasized.

REFERENCES

[1] Leino, T. (1976) Hyvän äänen spektripiirteitä. (In Finnish) (The spectral characteristics of good voice). Licenciate Thesis in Logopedics, Department of Phonetics. University of Helsinki.

[2] Leino, T. (1994) Long-term average spectrum study on speaking voice quality in male actors. In *SMAC93 Proceedings of the Stockholm Music Acoustics Conference July 28 - August 1, 1993*, A. Friberg, J. Iwarsson, E. Jansson & J. Sundberg, eds., pp 206-210.

[3] Gauffin, J. & Sundberg, J. (1980) Data on the glottal voice source behaviour in vowel production. *STL-QPSR (Speech Transmission Laboratory, Quarterly Progress and Status Report)*, Royal Institute of Technology, Stockholm, 2-3, 61-70.

[4] Frøkjær-Jensen, B., Prytz, S. (1976) Registration of voice quality. *Brüel & Kjær Technical Review* 3, 3-17.

[5] Wedin, S., Leanderson, R. & Wedin, L. (1977) Evaluation of voice training, In *Proceedings of the International Association of Logopedics and Phoniatrics, Ed. N.H.Buch, Copenhagen 1977*, pp. 361-381.

[6] Blomberg, M. & Elenius, K. (1970) Statistical analysis of speech signals.*STL-QPSR* 4, 1-8.

[7] Nolan, F. (1983) *The phonetic bases of speaker recognition.* Cambridge: Cambridge University Press.

[8] Sundberg, J. (1974) Articulatory interpretation of the singing formant. *Journal of the Acoustical Society of America* 55, 838-844.