TR－A－0099

Acoustic and Physiological Characteristics of Traditional Singing in Japan

Noriko KOBAYASHI, Yoh'ichi TOHKURA, Seiichi TENPAKU and Seiji NIIMI

1991.1.23

ATR視聴覚機構研究所

〒619-02 京都府相楽郡精華町乾谷 ☎07749-5-1411

ATR Auditory and Visual Perception Research Laboratories

Inuidani, Seika-cho, Soraku-gun, Kyoto 619-02 Japan

Telephone: +81-7749-5-1411
Facsimile: +81-7749-5-1408
Telex: 5452-516 ATR J

ACOUSTIC AND PHYSIOLOGICAL CHARACTERISTICS OF TRADITIONAL SINGING IN JAPAN*

NORIKO KOBAYASHI, YOH'ICHI TOHKURA, SEIICHI TENPAKU
ATR Auditory and Visual Perception Research Laboratories
and
SEIJI NIIMI
Faculty of Medicine, University of Tokyo

ABSTRACT

The purpose of this study was twofold: to understand the vocal differences in the many styles of singing found in Japanese culture and to expand our knowledge of human voice production beyond that found in Western classical singing. Four professional singers of traditional Japanese singing modes (Sohkyoku, Yohkyoku, and Min-yoh) were the subjects in our study. Recordings of the audio signal and video-fiberoptic images of the singers' larynges were made during singing. Prominently high energy concentration near 4-5 kHz were found in the singing voice, which could be considered as a "Japanese singing formant". Various acoustic patterns in terms of the phase relationships between F0 and power were found in the production of tone ornaments. Dynamic and continuous movements of the larynx were notable in Japanese singing when compared with Western classical singing, in which the vocal tract looks to be kept relatively stable.

*This paper has been submitted to Music Perception for publication.

INTRODUCTION

Numerous studies have been conducted on various aspects of singing. Generally theese studies assume a distinction between the singing voice and the speaking voice. However, what we call "singing voice" is not uni-modal since there exist a variety of singing modes or styles around the world. Different singing modes have evolved with different cultural traditions of a particular society where the singing mode has become a tradition. Their study and interpretation are susceptible to the aesthetic bias of the society that performs the investigation.

In the field of singing research to date, so-called "classical singing" as in *bel canto* has been of major interest. This classical singing mode developed in Western Europe between the Renaissance and the nineteenth century. Research on other modes of singing is scant even though they have been in existence for a longer period of time. Thus, one of the well-known acoustic findings for singing voice has been that of the "singing formant" found in male classical singing (Bartholomew, 1934; Vennard, 1967; Sundberg, 1973, 1974; Schutte and Miller, 1985). The "singing formant", a concentration of acoustic energy at about 3 kHz, has been found to have perceptual significance in good voice quality of classical singing (Winckel, 1952; Sundberg, 1973). However, it is not known whether singing modes in other cultures have specific acoustic characteristics equivalent to the Western classical "singing formant".

Like many other societies, the Japanese have a rich singing heritage. The voice quality of traditional Japanese singing may sound peculiar to Western ears, but not to the Japanese to whom it is familiar and beautiful. However, without specific perceptual data, it is not clear actually how traditional Japanese singing is perceived by listeners of other cultures. As Ishii (1962) pointed out, some Japanese singing modes with modal voice quality (*"jigoe"*) may sound painful and unhealthy to western classical singing teachers and voice clinicians even though the voice quality sounds natural to Japanese listeners. Based on observation of the larynx, Ishii (1962) described the Japanese modal voice as "throaty" since it appears to be produced with excessive tension associated with the adduction of the false folds- a configuration generally described as a hyperfunctional and unhealthy phonatory mechanism. In speech pathology, phonation with hyper-tension or hyper-adduction of the true and false vocal folds is generally considered to traumatize the vocal folds (Sonninen, Damste, Jol, and Fokkens, 1972; Boone, 1977; Prater and Swift, 1984). No finding has been reported so far to show clearly that traditional Japanese singing damages the voice. Furthermore, it is common for traditional Japanese singers to continue singing into old age without evidence of serious vocal dysfunction such as vocal nodules and chronic laryngitis. This suggests that these singers may develop special controls over the phonatory system that preserve vocal health no matter what the perception of Western classical voice experts is.

To begin the study of these "other ways" of using the voice, acoustic analyses and fiberoptic observation were made of the voice of three kinds of traditional Japanese singing.

METHOD

The singing modes examined in our study are Sohkyoku (singing accompanied by the "*koto*", an instrument with 13 strings), Yohkyoku (Noh recitation), and Min-yoh (folk song singing). Subjects were four professional singers--one from each of the three genres and one who has had professional careers in both Sohkyoku and Yohkyoku (Table. 1).

Recordings of the audio signal and the video-fiberoptic images of the singers' larynges were made while they were phonating. For the fiberoptic observation, topical anesthesia was applied to the subject's nasal cavity and a fiberscope (Pentax, FNL-10S) with 3.5 mm diameter was inserted through her/his nostril into the middle of the pharynx in order to view the larynx and the lower part of the pharynx. Subjects were instructed to sing a variety of phrases typical of their singing modes. No accompanying musical instrument was used during the recording sessions. Sustained phonation of vowels in speaking voice was also recorded. In order to assure that they were not phonating in singing voice when they were supposed to phonate in speaking voice, the singers' post-hoc subjective judgment on whether each sang or not was made for each speaking trial. Any sung samples were eliminated from the data based on both the singers' reports and the experimenters' perceptual impression during the experiments. For acoustic analyses, audio recordings of the same samples were made on a separate occasion in an anechoic room, using a digital audio tape-recorder (SONY, TCD-D10).

The recorded audio samples were re-digitized on to a computer using 16 bit quantization, a 32 kHz sampling rate, and a 16 kHz low-pass filter. Acoustic analyses for examining the spectral characteristics were made using 28th order LPC and 1024 point FFT methods, with 30 ms frame length and 40ms frame period. A Hanning window was used as the time window for the analysis. Fundamental frequency (F0) was measured by manually extracting the pitch periods through visual inspection of the acoustic waveforms. The pitch period was defined as the time between zero-crossings within one cycle. Values for power, represented in dB, were obtained pitch- synchronously by calculating mean square amplitude values within a pitch period.

RESULTS

1. Acoustic characteristics

Acoustic characteristics of the singing and speaking voice samples were compared for sustained and stable parts of vowel /e/ with nearly equivalent F0 values. In Fig. 1 both FFT and LPC derived spectra are shown for two types of phonation by a Sohkyoku singer (SH). The LPC derived spectra are not sufficiently reliable to extract accurate formant frequencies, but they are still useful for showing the overall spectral envelopes of the voice. Therefore, for spectral comparison of singing and speaking voice samples, both types of spectral analyses were used to determine the fine acoustic structure.

To extract predominant spectral peaks, the frequency range between 0 and 6.4 kHz was devided into three ranges; 0-2.3, 2.3-3.5, 3.5-6.4 kHz. The medium range (i.e., 2.3-3.5 kHz) was chosen based upon Kasuya's (1968) results which suggested that the formant frequencies of F2 and F3 for Japanese /e/ in female and child voices fall primarily within this range. Children's formant frequencies were taken into consideration because the singers in the present study were small. The physical height of the singers, for example, was only between 145 cm (4'8") and 154 cm (5'1").

In the speaking voice (top in Fig.1), the spectral envelope shows a decending pattern in which the FFT local peaks (indicated by arrows) within each of the three frequency ranges are gradually lowered as the frequency increases. In the singing voice (bottom in Fig. 1), on the other hand, the prominence of the highest peak of the FFT spectrum around 4 kHz is higher than that in the medium frequency range, even though the spectral slope is generally descending.

Similar patterns are found for the other Sohkyoku singer (NT) as seen in Fig.2. In the speaking voice (top in Fig.2), the highest spectral peak in the medium frequency range is almost the same as the peak around 4-4.5 kHz (indicated by an arrow); but, it is still slightly more prominent in the two singing voice samples. The voice samples of the Min-yoh singer (HK) also reveal similar patterns (Fig.3). For convenience, a spectral peak observed at 4-4.5 kHz in a spectral envelope will be called Fs hereafter.

To make a quantitative comparison between the speaking and singing voices, amplitude differences (in dB) between the highest FFT spectrum peak around the Fs and the one in the medium frequency range were calculated by simple subtraction (Fig. 4). Negative dB values on the ordinate in Fig.4 are obtained when the amplitude is higher for the medium frequency range than for the Fs, while positive dB values are obtained when the amplitude is higher for the Fs than for the medium frequency range. Thus, this figure shows the amplitude increase or decrease of the Fs relative to the peak in the medium frequency range for the speaking voice (white column) and the singing voice (grey column). As the vertical upward arrows in this figure show, the Fs has less amplitude decrease relative to the medium frequency range for the singing voice than for the speaking voice in all singing samples of SH and NT, and one of HK. For one singing voice sample of HK, the amplitude decrease of the Fs relative to the peak in the medium frequency range is greater than that of the speaking voice sample. Comparison of singing and speaking voices for a male Yohkyoku singer (HS) was not made since he tended to sing when he was supposed to produce sustained vowels in speaking voice.

Local changes of voice inserted in relatively smooth singing are frequently made for tone ornament in many modes of traditional Japanese singing. Figure 5(a) is a typical example of tone ornament in a Sohkyoku phrase sung by SH. Nine parts of tone ornament, indicated by the grey bars under the acoustic waveform, were produced in approximately nine seconds of singing. Acoustic analyses of these changes with regard to the relationship between F0 and power revealed that there were different patterns of tone ornament. Figure 5(b) shows a typical tone ornament "*atari*", which is perceptually characterized as a quick and brief change of voice. In the time period between 200 ms and 300 ms in this figure, a nearly reciprocal phase relationship between F0 and power is seen. Another example of phase reciprocity is found in the time period between 440 ms and 540 ms. Figure 5(c) is an example of another type of tone ornament "*furi*", which has a perceptual impression of slow and large fluctuations of voice. Slower changes of F0 and power with a reciprocal phase relationship can be observed in this figure.

For a vibrato-like voice fluctuation produced by the other Sohkyoku singer (NT), there is a synchronous phase relationship between F0 and power (Fig. 6). In this figure, an increase of F0 is accompanied by an increase of power, and a decrease of F0 is accompanied by a decrease of power. NT exhibited two kinds of acoustic patterns for "*atari*"; the reciprocal and synchronous phase relationships between F0 and power.

Figure 7 shows an example of tone ornament of Min-yoh, called "*kobushi*", produced by HK. Here, F0 is changed quickly; the initial decrease, the following increase, and the second decrease are achieved in about 200ms. Changes of power are made in shorter periods of time than those of F0, and there is no clear phase relationship between F0 and power.

2. Physiological characteristics

Fiberoptic observation revealed various kinds of dynamic behaviours in the singers' articulatory organs and larynges. In Sohkyoku, frequent changes in the vertical laryngeal position were associated mainly with the changes of pitch; the larynx was high for high pitch and low for low pitch. The false folds were not particularly adducted during relatively stable phonation as seen in a schematic drawing of the laryngeal view (Fig. 8), which is a tracing of a frame extracted from the video fiberoptic recording. No excessive adduction of the vocal folds or apparent abusive laryngeal configurations or gestures such as constriction of the false vocal folds were observed in this mode.

In Yohkyoku, the larynx appeared lowered relative to resting or breathing position immediately before singing started and the low position was maintained most of the time. As with Sohkyoku, no remarkable laryngeal constriction was observed during stable and sustained phonation in this mode (Fig. 9). The tongue and the epiglottis were vigorously in motion associated with articulation and the view of the vocal folds was occasionally obstructed particularly during articulation of back vowels.

In Min-yoh, the laryngeal position changed frequently as pitch changed. Unlike Sohkyoku or Yohkyoku, the false folds were remarkably adducted throughout singing particularly in the anterior parts (Fig. 10). The so-called aryepiglottic sphincter (the grey circle in the figure) looked slightly narrowed during phonation.

Various laryngeal movements were observed for the production of tone ornament in Sohkyoku, Yohkyoku, and Min-yoh. Slight adduction of the false folds was briefly observed in Sohkyoku sung by SH (Fig. 11) and NT as well as in Yohkyoku by HS. These two modes also exhibited a bending-like distortion of the vocal folds in the anterior-posterior dimension when producing tone ornament. Occasional excessive adduction or compression of the vocal folds was commonly observed in all modes.

DISCUSSION

The "singing formant" of Western classical singing is found in the acoustic spectrum of the professional singing voice as an energy concentration around 3 kHz. Sundberg (1973) suggested that the "singing formant" may play an important role in the perception of a "good voice", as voice with that particular formant can be heard over loud orchestra sounds. Using an acoustic model of the vocal tract, Sundberg (1974) further suggested that the "singing formant" is generated by expanding the lower part of the pharyngeal cavity and depressing the larynx. As the effects of the vocal tract configuration described above, posited Sundberg (1974), an additional cavity with the resonance frequency of approximately 3 kHz is created above the glottis and its resonance frequency is independent of the articulatory movements above it.

In our study, more prominent FFT spectral peaks in the frequency region of 4-5 kHz were found for the singing voice samples than for the speaking voice samples. These peaks may be considered as the "singing formant" of the traditional Japanese singing mode. The prominence of the peaks in the singing voice compared to that of the speaking voice, however, may not be so distinct as the one reported by Sundberg (1973, 1974) and other researchers of Western classical singing. It is mainly because the speaking voice of our subjects also sometimes has slightly enhanced energy concentration at 4-5 kHz but not so much as the singing voice. It should be noted here that the FFT spectral peak around Fs rather than strictly Fs itself has been chosen for comparison in our study since it was difficult to pinpoint the exact frequency value of Fs for some speech samples. Furthermore, as accurate formant extraction is difficult no matter which method is used, we considered it relatively more reliable and safe to identify the Fs by the method described above.

The spectral peak around Fs in traditional Japanese singing is much higher in frequency than the one at around 3 kHz found in classical singing of the Western culture. There seem to be two possible explanations for the difference between the two singing modes. One is perhaps the difference in the singer's vocal tract size. The Japanese singers in our study are short with the approximate height of 145-154 cm (4'8"-5'1") and it is assumed that their vocal tracts are much smaller than those of Western singers. Higher peak formant frequency is generated in the shorter vocal tract if we treat the vocal tract as a simple acoustic tube. The other possibility is that the singers who may or may not have small vocal tracts actively make some physiological adjustments in the phonatory or articulatory systems in order to generate a specific voice quality preferred in the Japanese culture. We have not yet found any evidence of the specific vocal tract configuration for the production of the "Japanese singing formant". It is quite likely, though, that the singers make some effort to maintain a specific feature of the vocal tract shape since the "singing formant" is preserved somehow in the vocal tracts which are continuously and dynamically changing during singing.In the next step of our study, we need to obtain data on multiple numbers of traditional Japanese singers with various vocal tract sizes to find whether the higher formant frequency value of the "Japanese singing formant" is generated mainly due to the small vocal tract size or due to special physiological adjustments. Also, data from singers whose singing repertoire covers both traditional Japanese singing and Western classical singing may provide us with some evidence to show how the "Japanese singing formant" is generated differently from the "Western singing formant" in the same vocal tract.

Various acoustic patterns of tone ornament were found in this study. Especially interesting is a reciprocal phase relationship between the fundamental frequency and power during the production of tone ornaments, such as "*atari*", "*furi*", and "*kobushi*". With regard to the phase relationship between F0 and power in vibrato of bel canto singing, it has been reported that the modulation patterns of the two parameters are in a synchronous phase (Seashore, 1932; Niimi, Horiguchi, Kobayashi, and Yamada ,1988) or they may proceed in a synchronous or asynchronous phase (Kwalwasser, 1926; Vennard, 1967; Large, 1979; Ramig and Shipp, 1986; Horii, 1989). Horii and Hata (1988) pointed out that the interaction between the resonance properties of the vocal tract and the sound source affects the acoustic modulation patterns of F0 and power. In our study, such tone ornament as seen in Fig. 5(b) and Fig. 7 demonstrate large excursions of F0 (200 Hz for "*atari*" and 110 Hz for "*kobushi*") relative to the change of power for less than 5 dB. The difference in the magnitude of the change for the two parameters is so significant that the modulation patterns do not seem to simply reflect a result of the source/resonance interaction. Rather, the modulation patterns observed in our study imply that the changes of F0 and power are made in the voice source, which is controlled by laryngeal and respiratory adjustments.

In most cases of speaking voice, F0 and power are known to have a synchronous relationship as the change in subglottal pressure affects both parameters such that the laryngeal control on F0 is kept constant. Therefore, it can be considered fairly "natural" for the physiological system to change F0 and power in a synchronous phase. A reciprocal phase relationship between F0 and power (i.e., increased F0 with decreased power) is commonly observed when voice quality is changed from modal to falsetto as F0 is increased. However, there is no such quality change based on the perceptual impression of the concerned tone samples in this study. Therefore, such an acoustic pattern does not seem to be simply a natural product of so-called register or quality change. It is not clear in this study how the physiological control of the variations of F0 and power are made especially when they have a reciprocal phase relationship. Is the larynx or subglottal pressure mainly responsible for the variations? Or, are the variations the acoustic product of the interaction between the source and transfer functions? As for bel canto singing, the controlling mechanism for tone fluctuation such as vibrato is not yet clear, either. Techniques such as inverse filtering and simultaneous recordings of subglottal pressure and electromyography of the laryngeal muscles may provide us with useful information about the possibility that the singers have the skills to control F0 and power independently against physiologically "natural and easy" ways of phonation, in which the two parameters vary with a synchronized phase relationship.

It is rather surprising that the fiberoptic image did not reveal particularly adducted false vocal folds in Sohkyoku and Yohkyoku despite the slightly harsh and strained impression of voice quality. Our finding is not in agreement with the "throaty voice" or a hyperfunctional phonation of Yohkyoku singing reported by Ishii (1962), who defines hyperfunction as phonation with excessive tension or constriction at the larynx and pharynx, based on the description by Brodnitz (1953). Physiological mechanisms for the production of strained voice quality without apparent constriction of the larynx are not yet clear. However, it seems likely that keeping the false folds abducted is an effective maneuver to avoid damage to the vocal folds particularly when various parts of the vocal tract are not stabilized but dynamically in motion. Although we can not find strong physiological evidence to support the above hypothesis, it is known that in phonation with excessive constriction of the larynx as seen in the case of creaky voice, for example, the vocal folds and the above portion in which the false folds are included are constricted (Daniloff, Schuckers, and Feth, 1980). Laryngeal tension, which is often associated with the constriction of the larynx has been known as a causal factor of vocal abuse (Aronson, 1980; Boone, 1977; Prater and Swift, 1984). It is presumed, therefore, that releasing tension in the false folds by opening them might prevent the vocal folds immediately below the false folds from having excessive tension. Detailed physiological research on the coordinated behaviour of the vocal folds and the false folds is needed to understand the significance of abduction and adduction of the false vocal folds.

The Min-yoh singer HK exhibited a slightly different fiberoptic image of the larynx when compared with the singers of the other singing modes. The adducted false folds observed throughout her singing could suggest that she sings with a lot of tension in the larynx and possibly constricts the vocal folds as well as the false vocal folds, taking the above mentioned view of vocal hygiene into account. As HK is the only Min-yoh singer in our study, it is not known whether such laryngeal behaviour is typical of Min-yoh singing or it is peculiar to her singing style.

Another characteristic finding about Min-yoh singing by HK is that the so-called aryepiglottic sphincter looks narrowed during singing. Garcia (1855) ascribed an important role of the epiglottis for the production of a brighter voice quality. Using a laryngeal mirror, he observed that the voice gains in brilliancy when the epiglottis was lowered and the orifice of the larynx was nearly closed. Based on the study with a rubber tube resonator model, Paget (1930) found that a nasal twang, which is usually perceived as a bright voice was produced by constriction above the glottis. Min-yoh, which is folk song singing and not so stylized or formal as Sohkyoku or Yohkyoku, is usually sung with a bright and loud voice. It is suggested that HK used a maneuver of constricting the aryepiglottic sphincter so that her voice quality could gain in brightness which is required for a "good" voice in Min-yoh singing. The relationship between vocal tract configuration and voice quality, however, needs further investigation.

Dynamic movements of the various parts in the vocal tract are one of the remarkable characteristics of traditional Japanese singing, while the vocal tract is relatively more stabilized in bel canto. Vertical movements of the larynx associated with the change of F0 have been considered to be a sign of poor or untrained singing in bel canto (Brodnitz, 1953). Voice quality in bel canto needs to be consistent in singing, and therefore, singers appear to keep the determined spectral characteristics rather constant (Schutte and Miller, 1985). This difference between bel canto singing and traditional Japanese singing may reflect the specific requirements for singing in the two different cultures. In Japan, various changes in the singing are required to express emotion, feeling, state, and atmosphere. As Kojima (1986) suggested, consistent phonation with beautiful and smooth voice in bel canto could bore some Japanese populations, who are used to delicate and frequent changes of voice.

In conclusion, our acoustic and physiological study of traditional Japanese singing revealed characteristic acoustic patterns and complicated behaviours of the singers' larynges. The obtained data are partly similar to Western classical singing but largely different from it. We believe that a "good" singing voice is not uni-modal and that the investigation of various singing modes can provide us with better understanding of both the nature of singing and physiological mechanisms for voice production and control.

REFERENCES

Aronson, A. *Clinical Voice Disorders: An Interdisciplinary Approach.* New York: Thieme-Stratton, Inc., 1980.

Bartholomew, W. T. A physical definition of 'good voice quality' in the male voice. *J. Acoust. Soc. Am.*, 1934, *23*:, 25-33.

Boone, D. R. *The Voice and Voice Therapy (Second Edition).* New Jersey: Prentice-Hall, Inc., 1977.

Brodnitz, F. S. *Keep Your Voice Healthy: A Guide to the Intelligent Use and Care of the Speaking and Singing Voice.* New York: Harper and Row Publisher, 1953.

Daniloff, R., Schuckers, G., and Feth, L. *The Physiology of Speech and Hearing.* New Jersey: Prentice-Hall, Inc., 1980.

Garcia, M. II. Observations on the human voice. *Proceedings of the Royal Society of London,* 1855, 399-410.

Horii, Y. Acoustic analysis of vocal vibrato: a theoretical interpretation of data. *Journal of Voice,* 1989, **3**, 36-43.

Horii, Y. and Hata, K. A note on phase relationships between frequency and amplitude modulations in vocal vibrato. *Folia Phoniatr,* 1988, ***40***, 303-311.

Ishii, S. *Koe no Shikumi (Mechanisms of Voice).* Tokyo: Ongakunotomo-sha, 1963.

Kasuya, H., Suzuki, H., and Kido, K. Changes in pitch and first three formant frequencies of five Japanese vowels with age and sex of speakers. *J. Acoust. Soc. Japan,* 1968, **24**, 355-365.

Kojima, H. *Yohkyoku no Ongakuteki Tokusei (Musical Features of Yohkyoku).* Tokyo: Ongakunotomo-sha, 1988.

Kwalwasser, J. The vibrato. *Psychol. Med. [Monogr Suppl],* 1926, ***36***, 84-108.

Large, J. An airflow study of vocal vibrato. In Weinberg, L. V. (Ed.), *Transcripts of the Eighth Symposium on Care of the Professional Voice.* New York: The Voice Foundation, 1979: Part I, 39-45.

Niimi, S., Horiguchi, S., and Kobayashi, N. Electromyographic study of vibrato and tremolo in singing. *Ann. Bull. RILP*, 1987, *21*, 153-164.

Paget, R. *Human Speech*. New York: Harcourt, Bruce and Co., 1930.

Prater, R. J., and Swift, R. W. *Manual of Voice Therapy*. Boston: Little, Brown and Company, 1984.

Ramig, L. A., and Shipp, T. Comparative measures of vocal tremor and vocal vibrato. Paper presented at the Fifteenth Symposium on Care of the Professional Voice, New York, June, 1986.

Schutte, H. K., and Miller, R. Individual parameters of the singer's formant. *Folia Phoniatr*, 1985, *37*, 31-35.

Seashore, C. E. *The Vibrato*. Iowa: University of Iowa Press, 1932.

Sonninen, A., Damste, P. H., Jol, J., and Fokkens, J. On vocal strain. *Folia Phoniatr.*, 1972, *24*, 321-336.

Sundberg, J. The source spectrum in professional singing, *Folia Phoniatr.*, 1973, *25*, 71-90.

Sundberg, J. Articulatory interpretation of the "singing formant". *J. Acoust. Soc. Am.*, 1973, *55*, 838-844.

Vennard, W. *Singing: the Mechanism and the Technique*. New York: Fischer, 1967.

Winckel, F. Electroakustische untersuchungen an der menschlichen stimme. *Folia Phoniatr.*, 1952, *4*, 93-113.

Table 1. Subjects and their singing styles.

SUBJECTS (SINGERS)

Name	Age	Singing style
SH	77	Sohkyoku (singing along "koto") Yohkyoku (Noh recitation)
NT	41	Sohkyoku
HS	62	Yohkyoku
HK	38	Min-yoh (folk song singing)

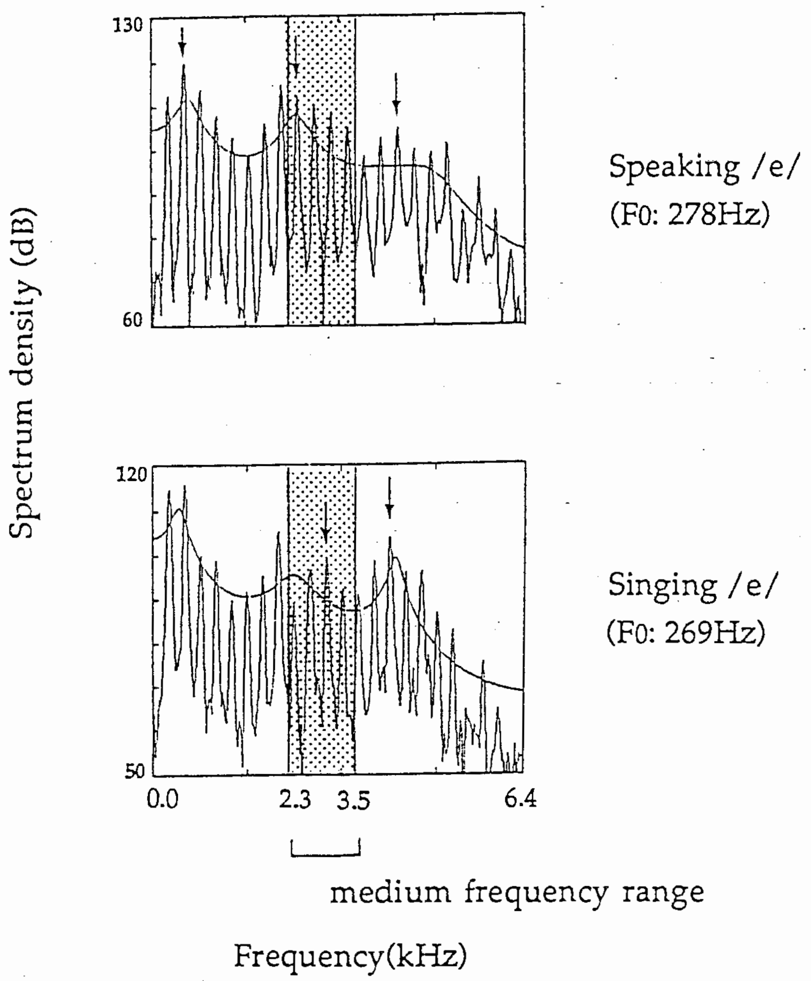

Fig.1. Spectrum slices of voice samples produced by a Sohkyoku singer SH.

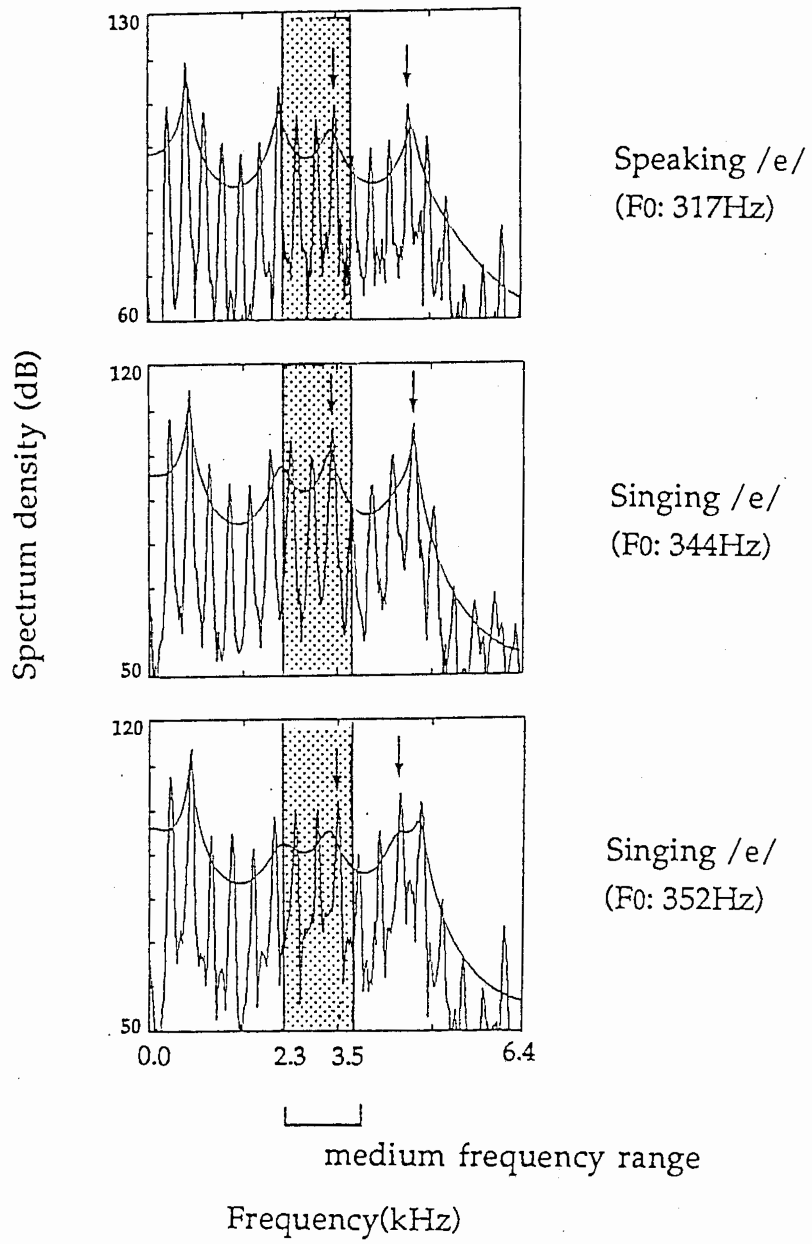

Fig.2. Spectrum slices of voice samples produced by a Sohkyoku singer NT.

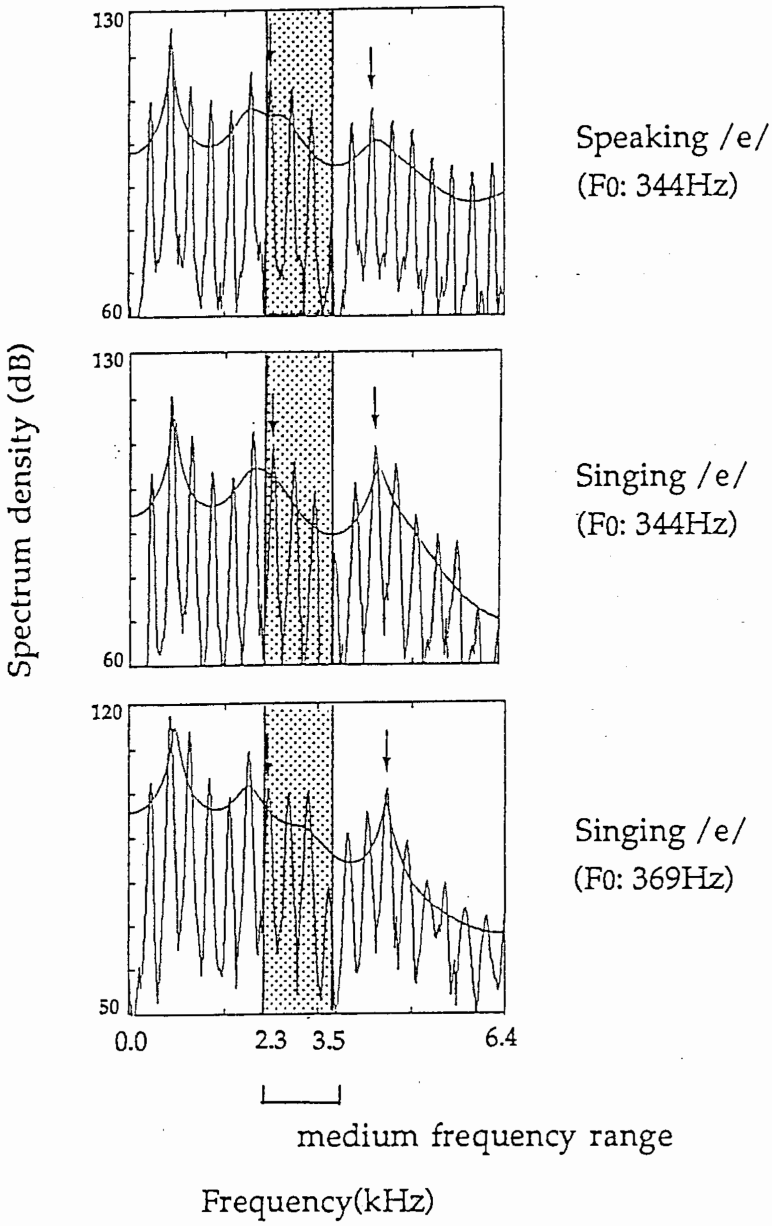

Fig.3. Spectrum slices of voice samples produced by a Min-yoh singer HK.

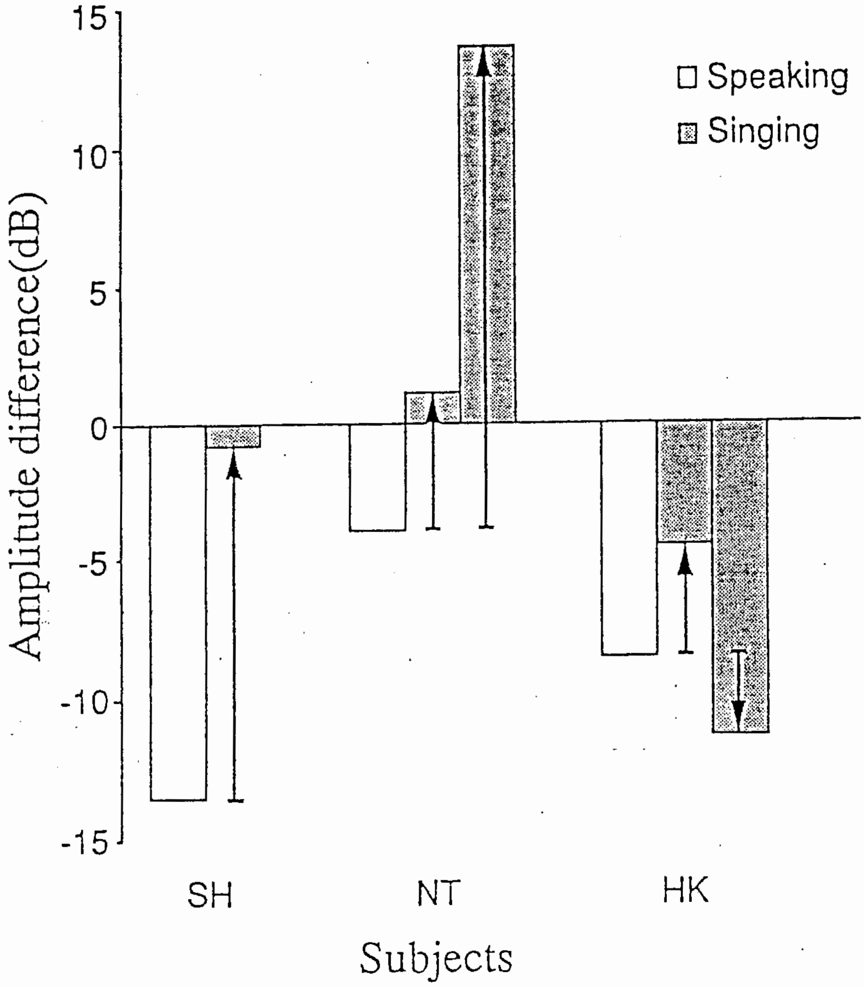

Fig.4. Amplitude differences between FFT spectral peaks of the Fs and those in the medium frequency range. The columns show the amplitude increase or decrease of the Fs relative to the peak of the medium frequency range for the speaking voice (white column) and the singing voice (grey column).

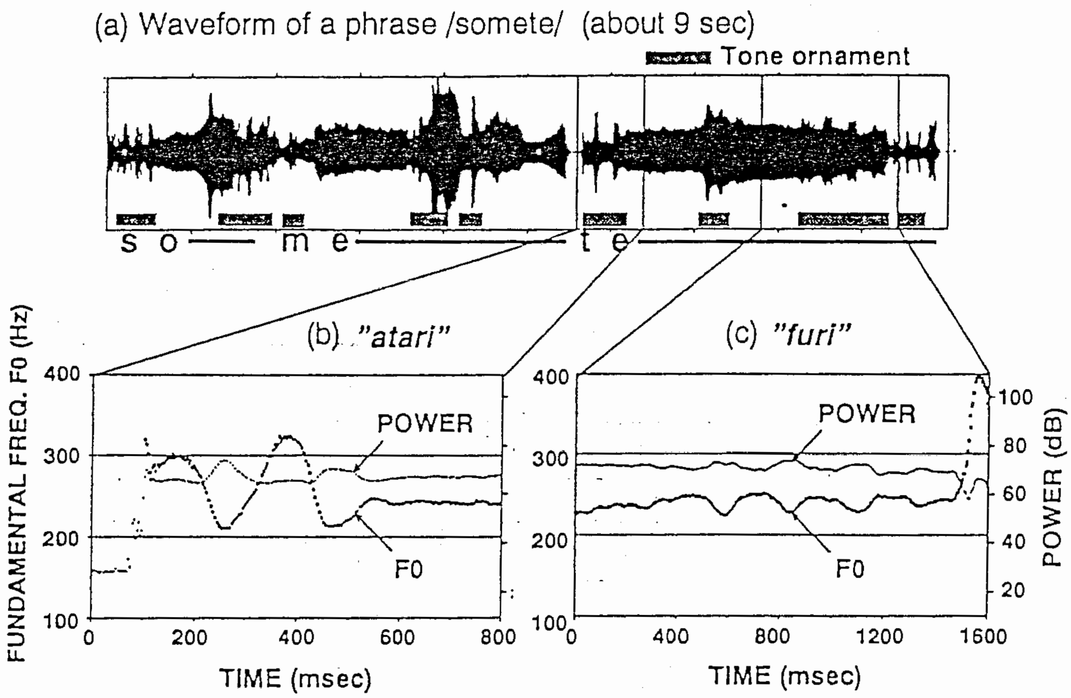

Fig.5. Acoustic waveforms of a Sohkyoku phrase produced by SH. Singing of /somete/ for about 9 seconds is shown in (a) with the grey horizontal bars representing parts of tone ornament. Changing patterns of power and F0 for two kinds of tone ornament "*atari*" and "*furi*" extracted from the upper figure (a) are shown in (b) and (c) respectively.

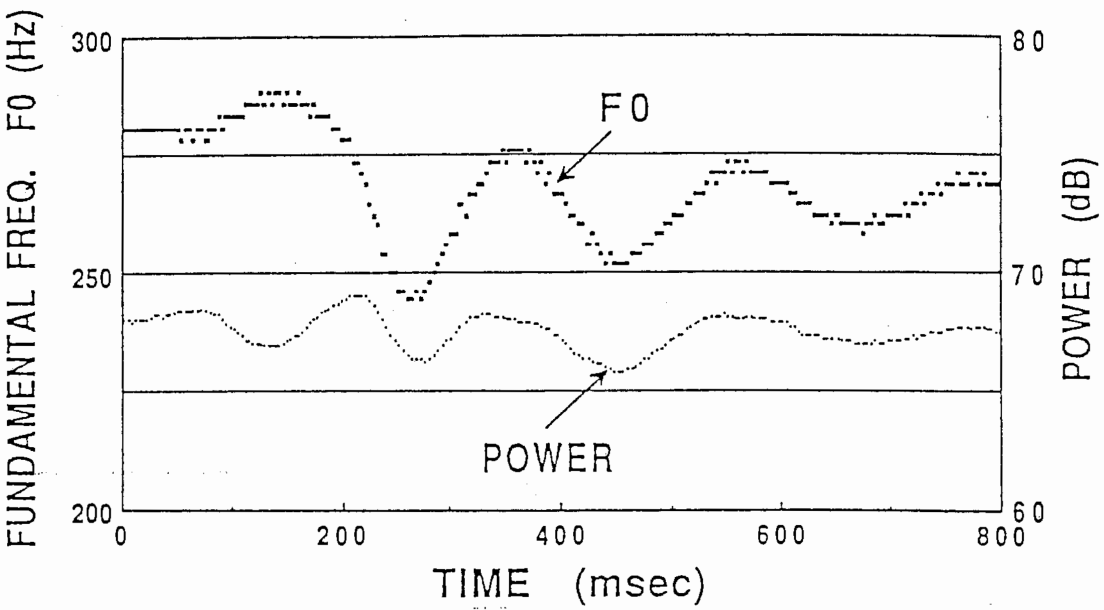

Fig. 6. Changing patterns of F0 and power for vibrato type of tone ornament in Sohkyoku singing by NT.

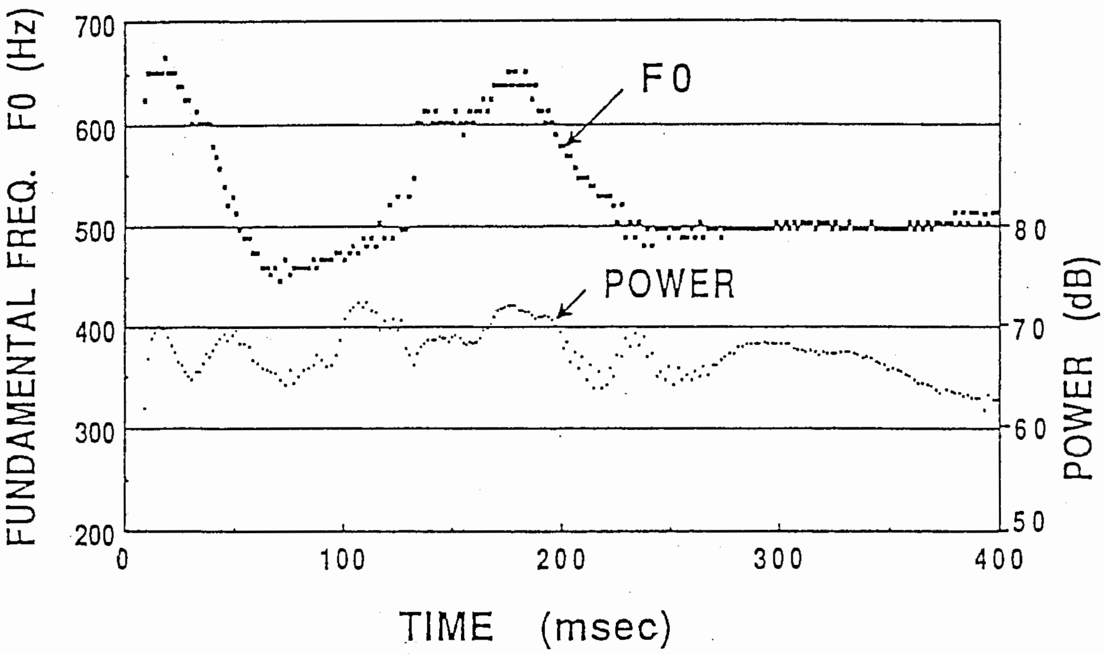

Fig. 7. Changing patterns of F0 and power for tone ornament "kobushi" in Min-yoh singing by HK.

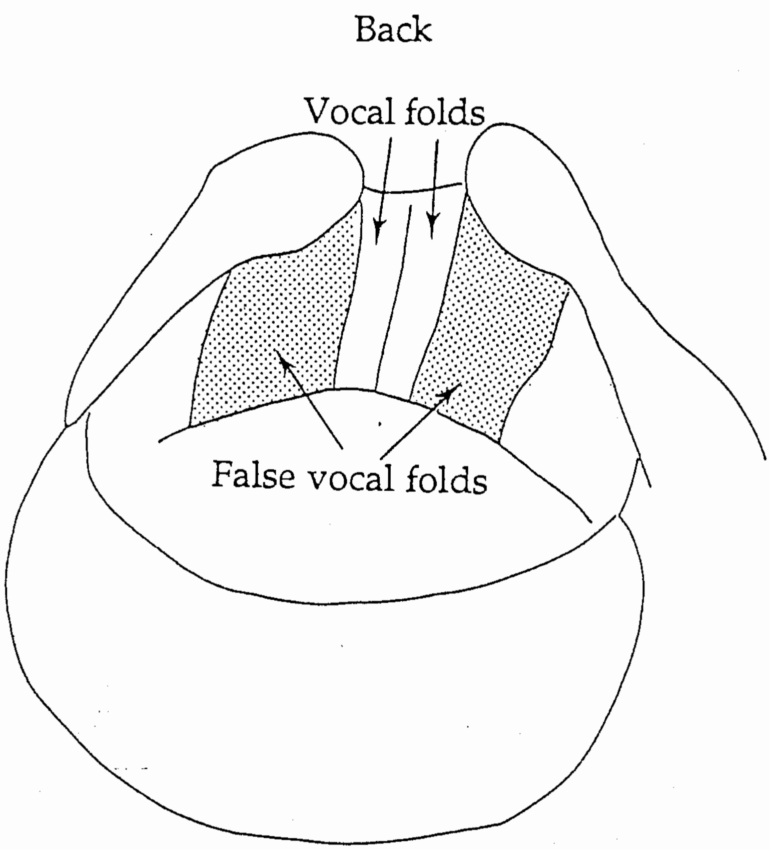

Fig.8. Fiberoptic view of the larynx during smooth singing of Sohkyoku by SH.

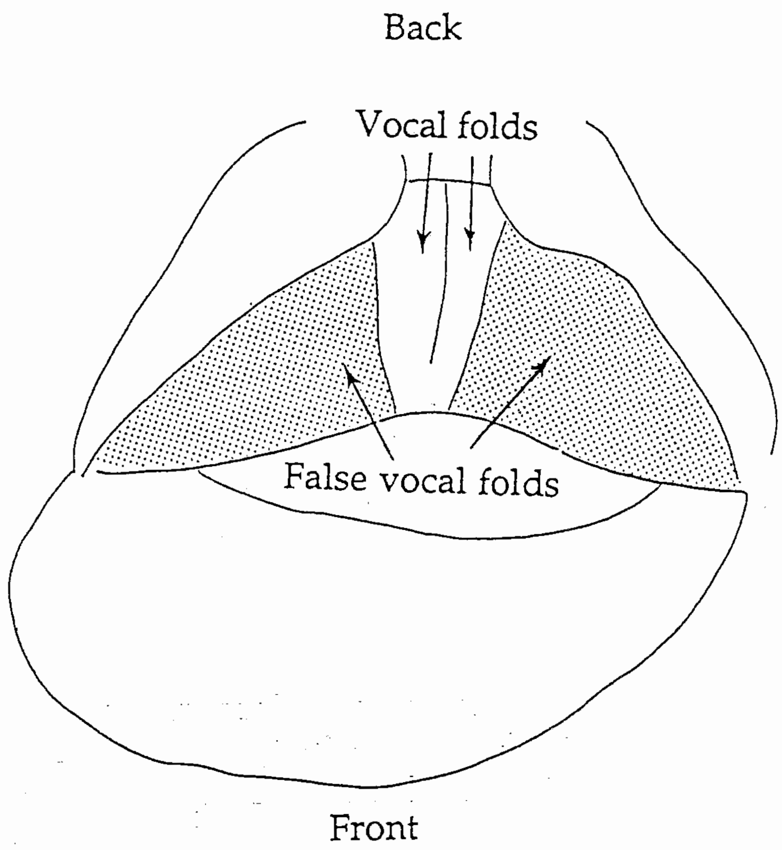

Fig. 9. Fiberoptic view of the larynx during smooth singing of Yohkyoku by HS.

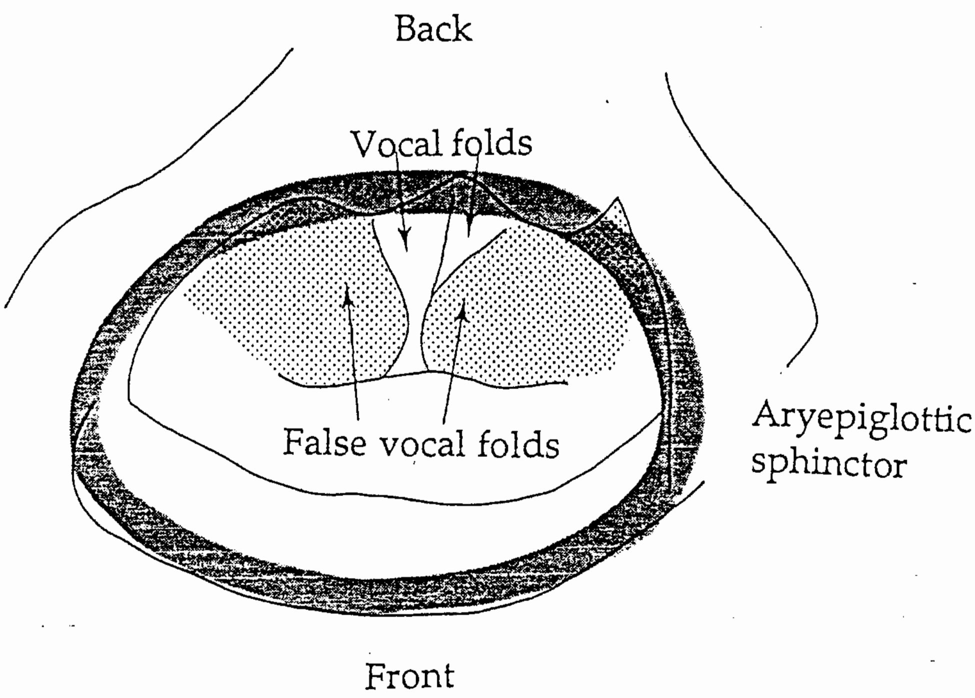

Fig. 10. Fiberoptic view of the larynx during Min-yoh singing by HK.

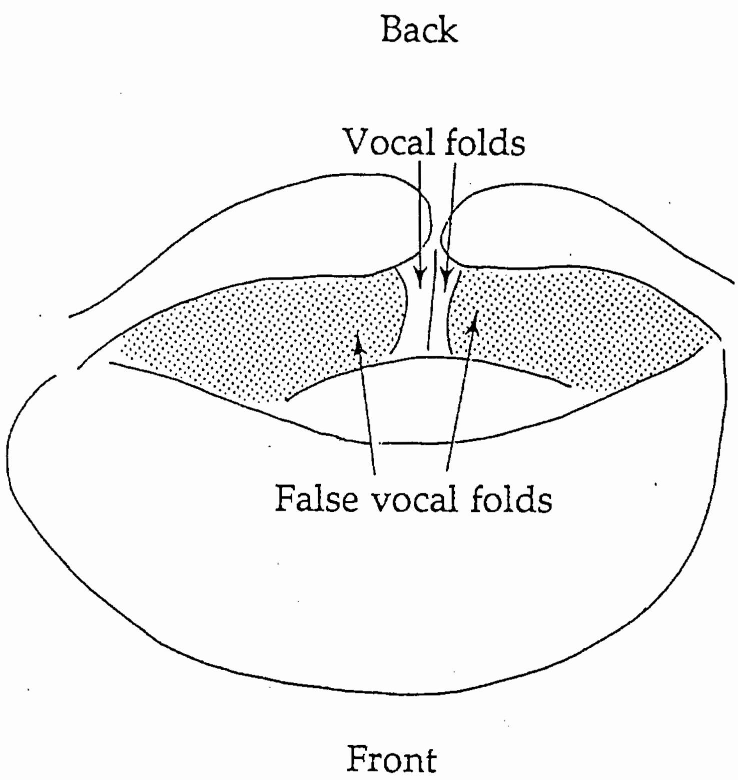

Fig. 11. Fiberoptic view of the larynx during the production of tone ornament in Sohkyoku by SH.