Sound Spectrographic Analysis of the Voice of Patients with Recurrent Larngeal Nerve Paralysis

Kwang-Moon Kim,* Yuki Kakita, Minoru Hirano

Department fo Otorhindarlngology, Institute of Logopedics and Phoniatrics,
*Yonsei University College of Mediune, Seoul, Korea**
Department of Otolaryngology, Kurume University, Kurume, Japan

Diagnostic evaluation for voice disorders requires a multidimensional approach : physiological examination of the larynx and the vocal tract, aerodynamic studies, examinations of vocal fold vibrations, psychoacoustic evaluations, acoustical analyses, and so on.

Previous studies with the use of a computer system suggested that the acoustic evaluation is quite promising for differentiating some causative diseases of voice disorders(Hiki et al., 1975a, b, 1976 ; Kakita et al., 1980). However, a well-equipped computer system for acoustic analysis is very expensive and not available in most voice clinics. A sound spectrograph is available in many voice clinics.

Recently, Imaizumi et al.(1980) repotted that acoustic parameters obtained from sound spectrograms were useful in differentiating pathological voices from normal voices. The purpose of this paper is to investigate the significance of acoustic parameters extracted from sound spectrograms in evaluating the voice of patients with recurrent laryngeal nerve paralysis.

Material and Method

Voice samples of 10 patients(5 males and 5 females) with recurrent laryngeal nerve paralysis were used. They were randomly selected from the audiolibrary at the voice laboratory of the Department of Otolaryngology, Kurume University Hospital. Age ranged from 19 to 85 years. As controls voice samples of 5 normal male and 5 normal female speakers, whose ages ranged from 20 to 44 years, were used.

The voice samples were recorded in the following way : The subjects uttered the five Japanese vowels/u/, /o/,/a/,/e/ and/i/at a comfortable pitch and loudness level. Each vowel was produced six times : three times with a long duration(approximately 3s) and three times with a short duration(approximately ls). The best sample among the three repetitions was selected later for acoustic analysis. An electret condenser microphone(Sony ECM-23) was placed at 20cm from the subject's mouth. It was connected to a tape recorder(TEAC A-6100 MK II). The tape speed was 19cm/s.

The recorded voice samples were analyzed using a sound spectrograph(Rion SG-07). In the present investigation, we studied the vowel /e/ because the report by Imaizumi et al(1980). was based on analyses of the vowel /e/ and we basically adopted their technique. The long samples were used for investigating the characteristics of the steady portion of phonation. The steadiest portion of 1 s duration was subjected to this analysis. The short samples were used for evaluating the characteristics of voice onset and offset.

The following nine acoustic parameters were obtained from spcetrograms :

1) Extent of fundamental frequency fluctuation. This was measured on a narrow-band filtered pattern. The extent of fluctuation was defined as the percent score of the ratio of the peak-to-peak value

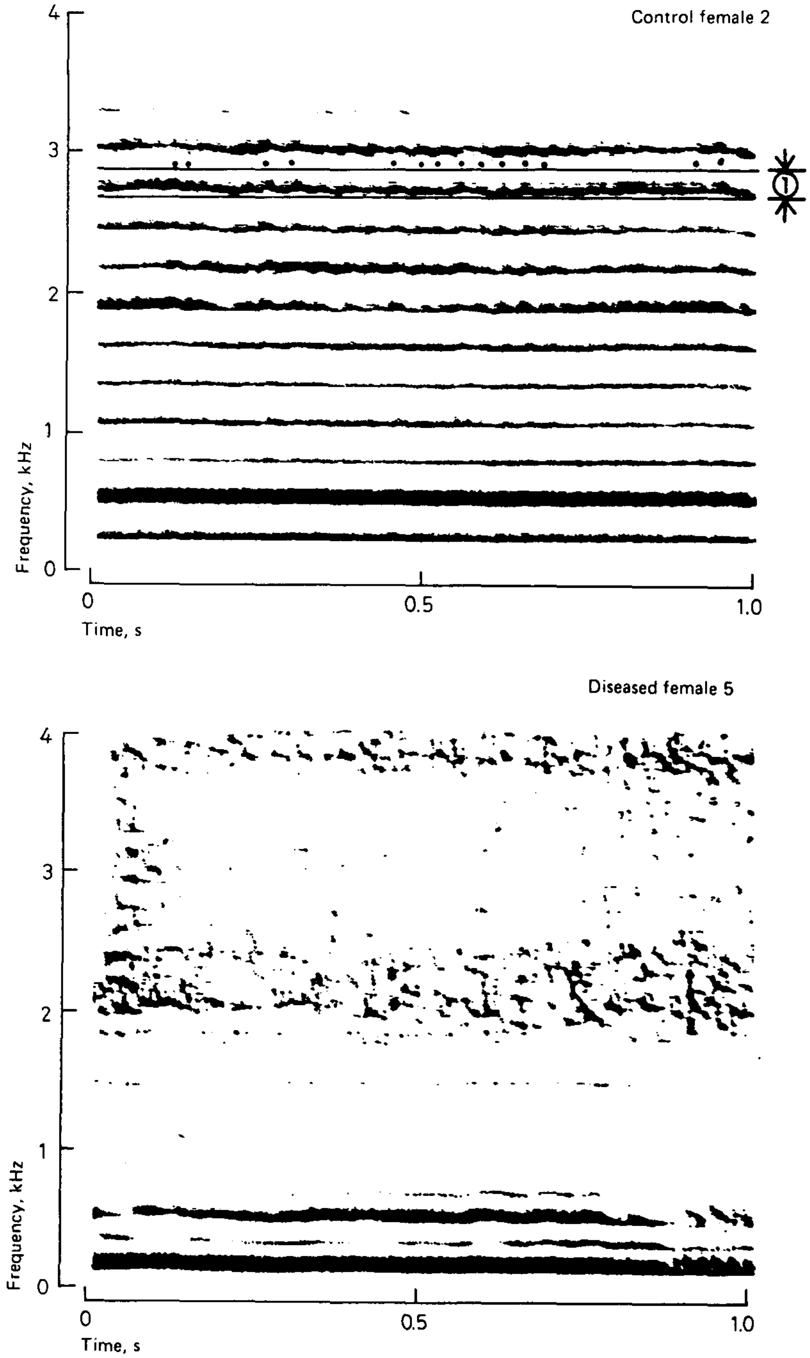

Fig. 1. Examples of the measurement of extent (1) and speed (2) of fundamental frequency fluctuation using a narrow-band filtered pattern. Both parameters are unmeasurable in the diseased subject. Figures in circles show parameter number. Dots indicate the counted peaks of fluctuation.

of fluctuation(Δ Fo) to the mean fundamental frequency($\bar{F}$ o). The actual measurements were made with a harmonic component as high as possible for the sake of accuracy(Fig. 1).

2) Speed of fundamental frequency fluctuation. This was defined as the number of positive peaks

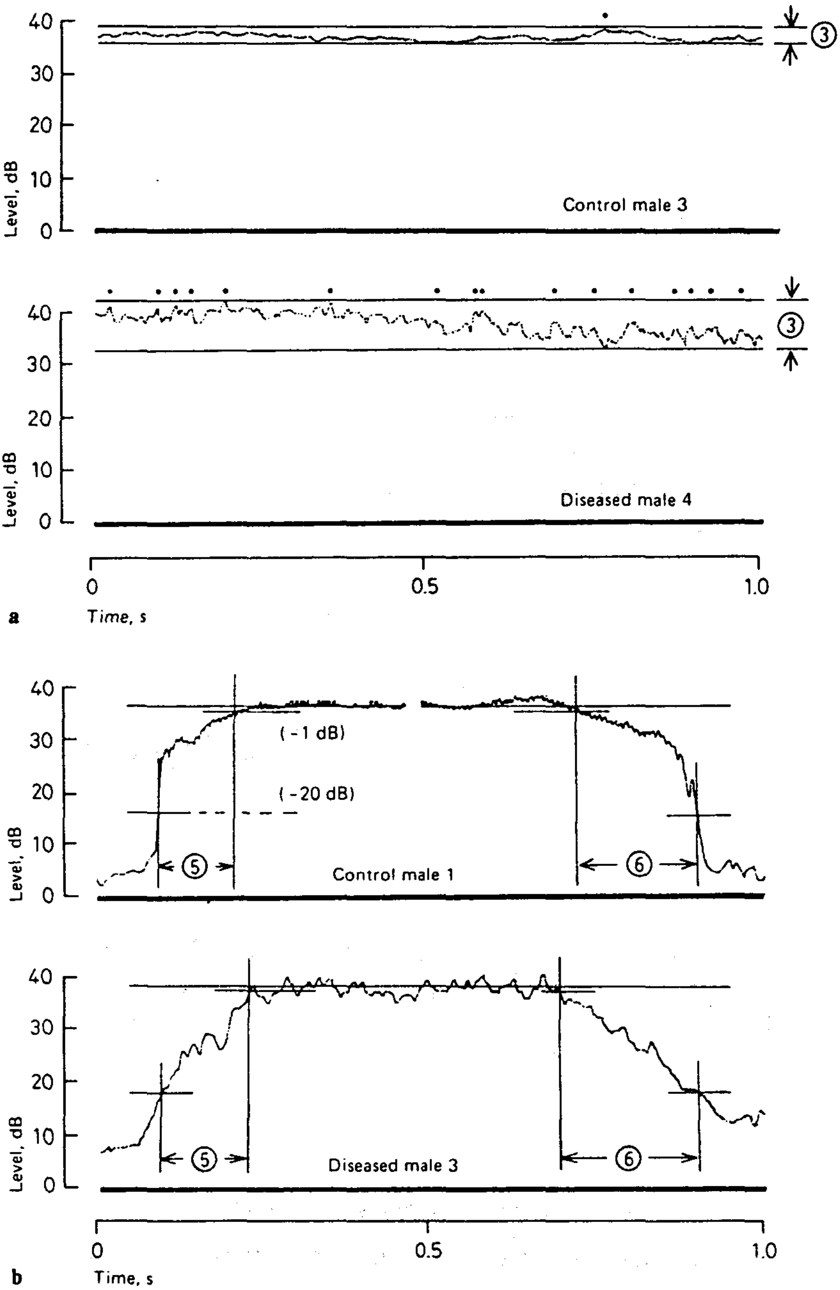

Fig. 2. Examples of the measurement of extent (3) and speed (4) of amplitude fluctuation and rising (5) and falling time (6) of amplitude using an amplitude display. a Extent and speed of amplitude fluctuation. b Rising and falling time.

within 1 s. For the measurement of this parameter, we used the 10th harmonics for both males and females. A detectable threshold of 1mm was applied to male subjects and 2mm to female subjects. The different threshold levels between the sexes were used because the fundamental frequency of females is roughly twice as high as that of males(Fig. 1).

3) Extent of amplitude fluctuation. This was defined as the peak-to-peak value in decibel measured on an amplitude display(Fig. 2a).

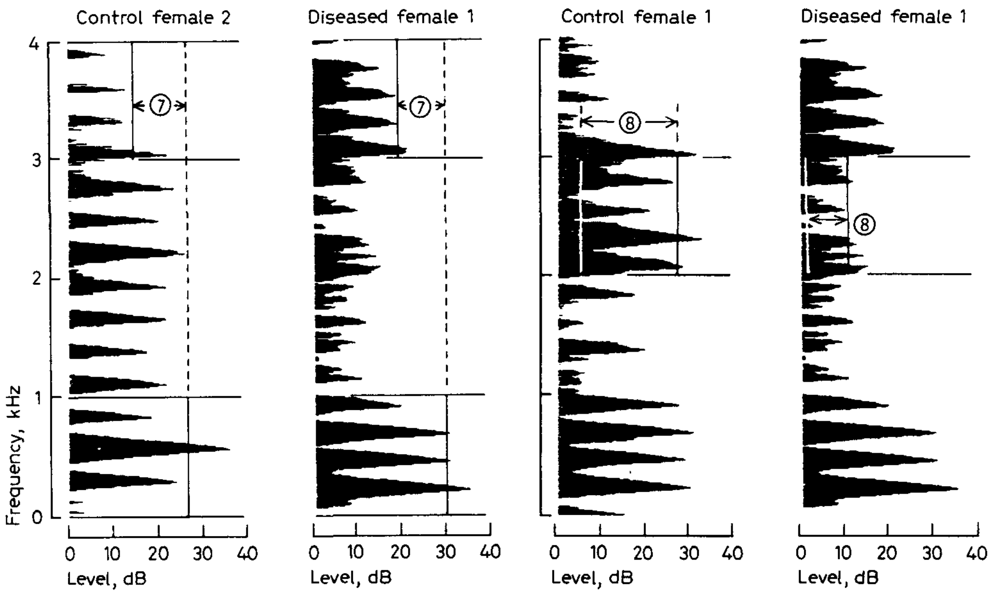

Fig. 3. Examples of the measurement of relative level of higher harmonic components (7) and relative level of noise (8) using a section display. a Relative level of higher harmonic components. b Relative level of noise.

4) Speed of amplitude fluctuation. This was defined as the number of positive peaks on an amplitude display within 1 s. Peaks of 3dB or greater from the adjacent troughs were counted(Fig. 2a).

5) Rising time and 6) falling time. These two parameters were measured on the amplitude display of the short vowel samples. The rising time was defined as the time required for an increase in the amplitude from 10%(−20dB SPL) to 90%(about −1dB SPL) of the steady level. The falling time was defined as the time span required for the decrease from 90 to 10% of the steady portion(Fig. 2b).

7) Relative level of higher harmonic components. A section display was used for the measurement of this parameter. This parameter was defined as the ratio of the intensity level between 3 and 4kHz to that below 1 kHz. In practice we first obtained two mean values : the value averaged over the peaks between 3 and 4kHz and that averaged over the peaks below 1kHz. The difference of these two mean values was adopted as the parameter showing the relative level of higher harmonic components(Fig. 3a).

8) Relative level of noise. This was defined as the ratio of the noise level to the level of the harmonic component and was estimated on a section display. In practice we first obtained two mean values between 2 and 3kHz : the value averaged over the peaks and that averaged over the troughs. The difference of these two mean values was employed as the parameter indicating the relative level of noise(Fig. 3b).

9) First formant frequency. This was estimated on a wide-band filtered pattern by careful inspection.

The first eight parameters are basically the same as those which Imaizumi et al(1980). described. In the present investigation, however, some of them were obtained in different ways : (a) In this investigation a threshold value was used for defining the peaks for parameters 2 and 4, while the peaks were determined by visual inspection in the work of Imaizumi et al(1980). (b) In this investigation the relative level for parameters 7 and 8 was determined on the basis of the mean values which were obtained by actual calculation. Imaizumi et al(1980). adopted visual inspection instead of actual calculation parameters. Parameter 9 was adopted on the basis of the work by Choi et al(1980). They reported that there is a significant decrease in formant frequencies in many pathological cases compared with normal cased.

Results

Measured values of each parameter are presented in tables I-V. In these tables, the subjects are divided into two major groups : control and diseased. Each group is further subdivided into male and female groups. In discussing the results, we will take the following procedures for each parameter :

1) If there is no statistical difference in the values between males and females in each of the two main groups, the values of the male and female subgroups are pooled and the main groups compared.

2) When the difference in the values between males and females is statistically significant in either of the control or the diseased group, males and females are discussed separately.

We employed the 'two-sample rank test' in determining the statistical significance. A 5% level was adopted in the test. This test is preferable when the number of samples is small. Before the two-sample rank test, extreme values were excluded on the basis of the Smirnoff test.

Extent of Fundamental Frequency(Fo) Fluctuation(Table 1). In the control group, the mean value of the extent of fundamental frequency fluctuation was greater for males than for females. However, the difference was not statistically significant.

In the diseased group, the extent of Fo fluctuation was measurable in only 2 cases out of 10. In the other 8 cases no clear-cut harmonics were detected. Imaizumi et al(1980). Assigned a value over 10% to unmeasurable cases. If this is applied, the extent of fluctuation for the diseased group is significantly greater than that for the control.

Speed of Fundamental Frequency(Fo) Fluctuation(Table 1). The mean value of the speed of F_0 fluctuation was greater for females than for males in the control group, but the difference was not statistically significant. In the diseased group, the speed of F_0 fluctuation was unmeasurable in 8 cases. There is no reasonable approximation for the unmeasurable cases. Therefore, statistical comparisons were not available with respect to the speed of F_0 fluctuation. We felt that the measurement of the speed of F_0 fluctuation was a difficult task even in the measurable cases.

Table 1. Extent and speed of fundamental frequency fluctuation

Case No.	Extent, %		Speed, n	
	Control	Diseased	Control	Diseased
Male				
1	4.1	UM	0	UM
2	3.6	7.9	0	28
3	4.8	UM	24[a]	UM
4	2.9	UM	7	UM
5	3.6	UM	17	UM
Mean± SD	3.80± 1.70	–	6.0± 8.0	–
Female				
1	3.1	UM	13	UM
2	2.7	UM	14	UM
3	2.7	UM	7	UM
4	1.6	3.6	6	27
5	3.2	UM	20	UM
Mean± SD	2.66± 0.63	–	12.0± 5.7	–

UM=Unmeasurable. Speed was defined as the number of positive peaks in 1 s.

[a] Value excluded by Smirnoff test. Mean and SD were calculated after the exclusion.

It is interesting to note that, in the 2 cases in which the extent and speed of F_0 fluctuation were measurable, the paretic vocal fold was fixed in the median position, whereas the paretic fold was fixed in the paramedian or the intermediate position in the other 8 cases.

Extent of Amplitude Fluctuation(Table 2). In both the control and diseased groups, there was no statistically significant difference in the extent of amplitude fluctuation between males and females. The mean value of the extent of the amplitude fluctuation was greater in the diseased group than in the control, but the difference was not significant statistically.

Speed of Amplitude Fluctuation(Table 2). There was no statistically significant difference in the speed of amplitude fluctuation in the control group. In the

Table 2. Extent and speed of amplitude fluctuation

Case No.	Extent, dB		Speed, n	
	Control	Diseased	Control	Diseased
Male				
1	4	12	0	21
2	4	5	1	0[a]
3	3	6	1	20
4	6	9	1	16
5	4	7	0	22
Mean± SD	4.2± 1.1	7.8± 2.8	0.6± 0.5	19.8± 2.6
Female				
1	6	4	1	0
2	7	5	3	0
3	5	2	0	0
4	5	11	3	6
5	6	10	6	6
Mean± SD	5.8± 0.8	6.4± 3.9	2.6± 2.3	2.4± 3.3

[a] Value excluded by Smirnoff test. Mean and SD were calculated after the exclusion.

Table 3. Rising and falling time

Case No.	Rising time, ms		Falling time, ms	
	Control	Diseased	Control	Diseased
Male				
1	119.0	142.9	182.5	111.1
2	71.4	55.6	55.6	95.2
3	91.3	135.0	87.3	206.3
4	146.8	182.5	15.9	87.3
5	119.0	99.2	103.2	190.5
Mean± SD	109.50± 28.96	123.04± 47.93	88.90± 62.06	138.08± 56. 0
Female				
1	23.8	67.5	79.4	166.7
2	31.7	63.5	71.4	87.3
3	23.8	71.4	206.3[a]	95.2
4	63.5	39.7[a]	87.3	127
5	43.7	71.4	123.0	63.5
Mean± SD	37.30± 16.76	68.45± 3.78	90.28± 22.76	107.94± 39.94
(M+F) mean± SD			89.51± 46.05	123.01± 48.53

[a] Value excluded by Smirnoff test. Mean and SD were calculated after the exclusion.

Table 4. Relative level of higher harmonic components and that of noise

Case No.	Harmonic components,dB		Noise, dB	
	Control	Diseased	Control	Diseased
Male				
1	− 7.8	− 7.2	− 9.5	− 7.2
2	+ 2.4	− 7.4	−17.9	− 7.8
3	−14.0	− 3.1	− 9.9	− 8.1
4	−10.2	− 7.2	−13.0	− 7.3
5	− 7.2	−11.4	−15.5	− 9.8
Mean± SD	− 7.36± 2.94	− 7.26± 2.94	−13.16± 3.61	− 8.04± 1.05
Female				
1	−17.1	−10.8	−21.7	− 9.0
2	+12.1	− 6.3	−23.7	− 8.1
3	−17.1	− 6.1	−21.2	−10.3
4	−20.1	−23.0[a]	−12.8	−14.2[a]
5	−13.1	− 8.3	− 8.1	− 9.5
Mean± SD	−15.90± 3.27	− 7.88± 2.19	−17.50002/6.71	− 9.23± 0.99

[a] Value excluded by Smirnoff test. Mean and SD were calculated after the exclusion.

diseased group, however, the speed of amplitude fluctuation was significantly greater in males than in females. The difference between the control and diseased group was statistically significant in males (the diseased being greater), but not in females.

Rising Time of the Amplitude(Table 3). In the control group, the rising time was significantly longer for males than for females. No significant difference between the two sexes was found in the diseased group. The mean rising time of the diseased group was longer than that of the control group in both males and females. The difference was statistically significant only in females.

Falling Time of the Amplitude(Table 3). The falling time did not show any significant differences between males and females of between the control and diseased groups.

Relative Level of Higher Harmonic Components (Table 4). In both the control and diseased groups, the relative level of higher harmonic components did not differ significantly between males and females. The relative level of higher harmonic components was significantly greater in the diseased group than in the control.

Relative Level of Noise(Table 4). In both the control and diseased groups, there was no significant difference in the relative level of noise between males and females. The relative level of noise was significantly greater in the diseased group than in the control.

Table 5. First formant frequency

Case No.	Control, Hz	Diseased, Hz
Male		
1	540	450
2	470	370
3	520	470
4	490	260
5	580	260
Mean± SD	520± 43	366± 96
Female		
1	560	490
2	540	490
3	560	430
4	640	510
5	450	270
Mean± SD	550± 68	458± 58
(M+F)mean± SD	535± 56	412± 89

First Formant Frequency(F_1)(Table 5). There was no significant difference in F_1 between males and females. The mean value of F_1 was smaller in the diseased group than in the control. The difference was statistically significant.

Discussion

Among the nine acoustic parameters investigated in this paper, significant differences between the control and diseased groups were found in the fluctuation of F_0 and relativel level of higher harmonic components, relative level of noise, and F_1. The fluctuation of amplitude may also differentiate the diseased group from the control. The large fluctuationof F_0 and amplitude in the cases of recurrent laryngeal nerve paralysis indicates that the vocal fold vibration is irregular. Irregular vibrations in paralytic cases have been observed with a stroboscope and a high-speed camera(Hirano et al., 1972, 1973 ; Schönhärl, 1960).

Our results revealed that irregular vibration is reflected on sound spectrographic displays. The high noise level in paralyzed cases indicates that the glottic closure is incomplete during phonation.

It is rather diffcult to explain why the level of higher harmonic components is higher in diseased cases. Further investigation is needed in this respect.

The tendency of low F_1 in the diseased group agrees with the results of Choi et al(1980). Provided that the vowel is the same(and especially for an open vowel like /a/ or /e/), the length of the vocal tract can affect F_1 most significantly among the formants, sinece F_1 is the lowest resonant frequency. Therefore, the the lowering of F_1 in the diseased group suggests that there can be a vocal tract lengthening. In paretic cases, there might be an abnormal muscular behavior during phonation which results in larynx lowering.

Résumé

Des analyses spectrographiques de la voix(voyelle /e/)ont été effectuées chez 10 malades présentant une paralysie du nerf récurrent et chez 10 sujets normaux. Neuf paramétres acoustiques ont été ainsi obtenus. L'importance de la fluctuation de la fréquence fondamentale, le niveau relatif des composants harmoniques les plus élevés, ainsi que le bruit et la fréquence du premier formant ont présenté les différences les plus significatives entre les malades et les sujets notmaux. La vitessed de fluctuation de l'smplitude et du temps de montée présente aussi des différences significatives, mais à un degrémoindre. La vitessed de fluctuation de la fréquence fondamentale, l'importance de la fluctuation de l'amplitude et le temps de retombée sont sans intérêt dans l'évaluation de la pathologie.

Zusammenfassung

Eine schallspektrographische Analyse der Stimme (Vokal /e/) wurde an 10 Patienten mit Paralyse des Nervus laryngeus recurrens und an 10 Gesunden durchgeführt. Neun akustische Parameter wurden aus den spektrographischen Daten abgeleitet. Die Grösse der Grundfrequenzschwankungen, das relative Niveau von harmonischen Komponenten und Lärm und die Frequenz des ersten Formanten zeigten die signifikantesten Unterschiede zwischen Patienten und Gesunden. Die Geschwindigkeit der Amplitudenschwankungen und die Anstiegszeit zeigten ebenfalls signifikante Unterschiede, aber in kleinerem Umfang. Die Geschwindigkeit der Grundfrequenzschwankungen, die Grösse der Amplitudenschwankungen und Abfallszeiten erwiesen sich als nicht brauchbar in der auswertenden Pathologie.

References

1) Choi, SE, Kim , GR : *The medico-sona-graphic study of Korean hoarseness dut to laryngeal pathology. Yonsei J. med. Sci., Seoul 13 : 82, 1980*

2) Hiki, S, Imaizumi S, Hirano, M, Matsushita, H, Kakita Y : *Acoustical analysis for voice disorders : an attempt towards clinical application. J Acoust Soc Am 58 : 111, 1975a*

3) Hiki S, Imaizumi S, Hirano M, Matsushita H, Kakita Y : *Acoustical analysis for voice disorders. Proc IEEE Int Conf Acoustics, Speech Signal Processing, p613, 1976*

4) Hiki S, Matsuoka K, Kakita Y, Imaizumi S, Hirano M, Matsushita H : *A study on acoustical analysis of hoarseness. J Axoust Soc Jap 31 : 504, 1975b*
5) Hirano M, Kawasaki H, Matsushita H, Yoshida Y, Koike Y : *Mode of vocal cord vibration in recurrent laryngeal nerve paralysis. Otolaryngol. Jap 76 : 721, 1973*
6) Hirano M, Nozoe I, Shin T, Maeyama T : *Vibration of the vocal cords in recurrent laryngeal nerve palsy. A stroboscopic investigation. Practica otol, Kyoto 65 : 1037, 1972*
7) Imaizumi S, Hiki S, Hirano M, Matsushita H : *Analysis of pathological voices with a sound spectrograph. J Acoust Soc Jap 36 : 9, 1980*
8) Kakita Y, Hirano M, Imaizumi S, Hiki S : *Discrimination of pathological voice based on acoustical analysis. Jap J Med Electronics biol Eng 18 : 13, 1980*
10) Schönhärl, E : *Die Stroboskopie in der praktischen Laryngologie(Thieme, Stuttgart 1960)*

Received : August 9, 1981
Accepted : August 26, 1981

Dr. Kwang Moon Kim, Department of Otolaryngology, Yonsei University College of Medicine, Severance Hospital, CPO Box 8044, Seoul(Korea)