Objective Documented Study on Stipulated Target-Tone.

Test of Vocal Controllability with the Electromyography background mechanism.

Hartono Abdoerrachman[1], Satoshi Imaizumi[2], Seiji Niimi[2], Masanobu Kumada[2]

Abstrak

Penderita disfonia mempunyai kesulitan untuk berkomunikasi sehari-hari, karena adanya kesulitan mengucapkan kata-kata secara tepat pada nada suara normal. Atas dasar itu, maka telah dilakukan analisa akustik dan elektromiografi pada penderita kelainan suara neurologik dan pembicara normal sebagai pembanding. Keadaan tersebut ternyata disebabkan karena adanya peningkatan nilai parameter analisa suara serta adanya penurunan aktivitas potensial otot-otot Vocalis dan Crycothyroideus. Sebaliknya pengucapan suara dengan nada yang tinggi (high pitch/falsetto) ternyata secara bermakna dapat meningkatkan kemampuan penderita untuk mengontrol pengucapan kata-kata. Hasil tersebut menunjukan bahwa peningkatan kemampuan pengucapan kata-kata terjadi karena adanya perbaikan kualitas kontraksi otot-otot suara.

Abstract

A study of acoustic analysis and electromyography upon neuro-pathological condition of voice disturbance and normal speakers had been conducted. This study was based on the complaint that dysphonic patients are having difficulties on controlling their voice volitionally and the accuracy for daily verbal communication. The results indicated that voice disturbance was related to the increased parameter values extracted on voice analysis and the decreased of the potential action of the Vocalis and Crycothyroid muscles. However, significant changes of vocal controllability was obtained during the high pitch/falsetto voice of the stipulated target tone. The phenomenon indicate that when subjects produced stipulated target-tone, the vocal controllability were improved due to the betterment of the muscles action.

***Keywords** : Voice analysis, acoustic analysis, vocal controllability, electromyography.*

INTRODUCTION

Patients with either neurological or laryngeal disorders tend to complains of having difficulties in controlling their voices, flexibly and volitionally for daily verbal communication.[1,2] It is well accepted that the voice production, especially the speech production is influenced complexly by the constellation works of laryngeal intrinsic muscles. It is known that the Vocalis and the Cryco-thyroid muscles are the two laryngeal intrinsic muscles which are primarily responsible for the increasing vocal pitch.

Imaizumi *et al.*,[3,4] Abdoerrachman *et al.*,[1,5] Hirose *et al.*,[6] have reported their investigations concerning the analysis of vocal controllability in different pathologies of the larynx. However, they mostly used a single mode of pitch and amplitude in producing the voice samples. Since that the daily conversation is full of intonations, thus, it is important to assess the voice conditions with different mode of phonations, the low and high pitchs. Voice controllability is defined as the ability of subjects to keep the fundamental frequency and amplitude of voice as constant as possible when instructed to produce sustained vowel.

The aim of this investigation was to explore the background mechanism of voice alteration base on a given task to the subjects to increase their voices to reach the stipulated target-tone (falsetto voice) through an objective electro-myographic (EMG) recording of the potential muscle action and assessment of their the vocal controllability.

Our results demonstrated that the spasmodic dysphonia (SPD) subjects have a higher perturbation for comfortable pitch (SPD-low) as compared to the high pitch (SPD-high, falsetto) voices. Moreover, by

[1] *Department of Otorhinolaryngology, Faculty of Medicine, University of Indonesia, Jakarta, Indonesia*
[2] *Research Institute of Logopedics and Phoniatrics, Fac. of Medicine, Univ. of Tokyo.*

producing the stipulated high tone (high pitch, falseto), the perturbation value was decreased in accordance to a better vocal controllability. EMG analysis revealed that the improvement of vocal controllability was due to either by increased number of motor firing units and/or a higher firing rate in a single unit of the Vocalis and Crycothyroid muscles.

METHOD

Subjects for the study were 2 SPD patients and 2 normal / healthy voice volunteers which have not any history of pathology affecting phonation as the control group. The modal voice samples produced at most comfortable level and pitch by healthy speakers.

Voice samples collection were performed by recording sustained vowel /a/ produced by subjects as long as they were able to. SPD patients were requested to produce sustained vowel /a/ at comfortable level and pitch (SPD-Low), and followed by stipulated high pitch (SPD-High, falsetto), in fractional steps. Recordings were made using in a sound attenuated room with the DAT tape-recorder, with constant distance from microphone to mouth of 15 cm, following repeated patient's practice.

Voice samples of sustained /a/ were then digitized through a 16-bit analog to digital (A/D) converter at a sampling rate of 40 kHz and stored on a disk controlled by computer. A half second segment was extracted by excluding the initial and final portions from each sample. The token segment were chosen from the regular-like portion of phonation by looking at the F_o time series and the sonogram, confirmed by perceptual judgment. Saturated recording of sustained vowel, and also erratic and disrupted portions were excluded.

Detection of local maximum points of the voice were done by using the wave-form matching and peak picking method, which was proposed by Imaizumi *et al.*[7]. Local maximum points of the voice waveform is corresponding to the vocal excitation epochs of each glottal cycle. This were followed by determination of the two time series of $F_o(i)$ and $A(i)$, as the fundamental frequency and the maximum amplitude of i-th glottal period.

Moreover, cycle by cycle perturbation quotients were calculated and several voice properties were also been extracted. Pitch Perturbation Quotient (PPQ) and Amplitude Perturbation Quotient (APQ) represents the amount of fast fluctuations in pitch (inverse F_o) and amplitude. However, these parameters do not explicitly illustrate how fast the fluctuations are.[8]

The additive noise level (Noise Level) is the difference in deci-Bell (dB) between the energy of the non-harmonic and that of the harmonic components, within 1-4 kHz frequency range. The harmonics and non-harmonics components were extracted from voice waveform using a comb filtering method.[9] This parameters represents the magnitude of noise components, which may be responsible to "breathy" and "hoarse" voice quality.

The overall variability of F_o (F_oV) is the percentage of the standard deviation normalized by the average of $F_o(i)$. A logarithmic transformation is used for statistical analysis. This parameter represents an instability of F_o. The larger values indicates the more instability in controlling F_o.

Fast fourier transformation of the power spectra were used to calculate the slow (F_oS) and fast (F_oF) fluctuation of the energies in $F_o(i)$. Followed by the calculation of the energies in the frequency ranges between $0<f<16$ and $16<=f<$average $F_o/2$. F symbolized the frequency in Hertz (Hz). Finally the logarithmic transformed values are normalized by DC level. F_oS represents the magnitude of F_o fluctuation which is slower than 16 Hz, while FoF is that of faster than 16 Hz.

The Vocalis and Cryco-thyroid muscles activities for comfortable voice (SPD-Low) and high pitch (falsetto, SPD-High) were recorded through bipolar hooked wire electrodes. The second halfs EMG record were token as the sample. This regular-like phase was extracted by looking at the F_o time series, excluding the initial and the final portions. To obtained quantitative data of the token samples, the EMG data were then executed and rectified, integrated, smoothed using EMGPRO-program, and finally normalized by 0.5 second segment of sinusoidal wave-form calibration (300uV/kHz).[10]

Calculation of the analysis of variance (ANOVA) were done using the STATVIEW computers program.

RESULTS

An experimental study to analyze the vocal perturbations and the potential action of the Vocalis and Cryco-thyroid muscles had been conducted upon normal subjects and neuro-pathological condition of dysphonia. The feature of voice profile as observed in computers monitor was shown in Figure 1.

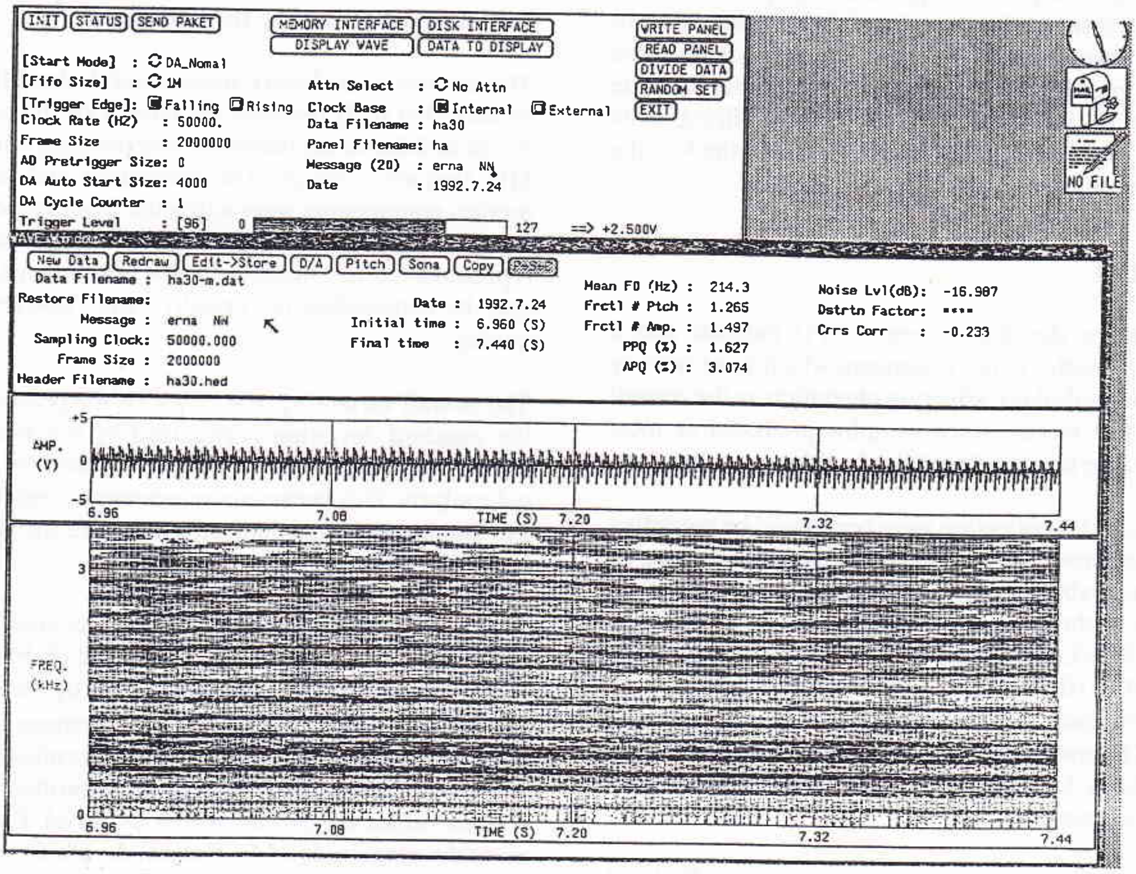

Figure 1. Feature of analyzed voice sample using "SONG"-program, with 4 panels: A) display of commands, B) display of extracted voice properties with the scores, C) acoustic wave-form, and D) voice spectrogram or the sonogram.

Voice perturbations which were revealed in several extracted parameters can be seen in Figure 3 (PPQ), Figure 4 (APQ), while Figures 5, 6 and 7 represented F_oV, F_oS and F_oF. Significant differences ($p<0.0001$) was noted on decrements between SPD-Low and SPD-High.

The simultaneous acoustic wave-form EMG pattern of the Vocalis and Cryco-thyroid muscles were shown in Figure 9. Figure 10 and 11 demonstrated the significant differences ($p<0.0001$) of the potential action during the SPD-Low and SPD-High.

DISCUSSION

SPD voice is acquainted as strain-strangle, striving, jerky, tremorous with intermittent voice stops. However, the "normal-like" pattern was still noted between the disruptive voice profile. This experiment was done with a special intention to these pattern.

The SPD-High (falsetto) speakers showed an increased mean value which exceeding the normal fundamental frequency (Figure 2). The increase of pitch was effected by the contraction and increased activation of The Cryco-thyroid and Vocalis muscles, which act as the adductor and tensor of the vocal folds.

The pitch perturbation quotient (PPQ), the amplitude perturbation quotient (APQ) and the fast and slow F_o perturbation (F_oF and F_oS) which were decreased significantly on SPD-High (falsetto) as compared to the SPD-Low speakers, indicated that by increasing the pitch into falsetto-tone there were a decreased in the magnitude scores of perturbations.

The overall variability of fundamental frequency (F_oV), which was represented the F_o instability found to be decreased significantly ($p<0.0001$) on SPD-High (falsetto voice). This data showing that the high pitch voices were less fluctuating, and more stable as com-

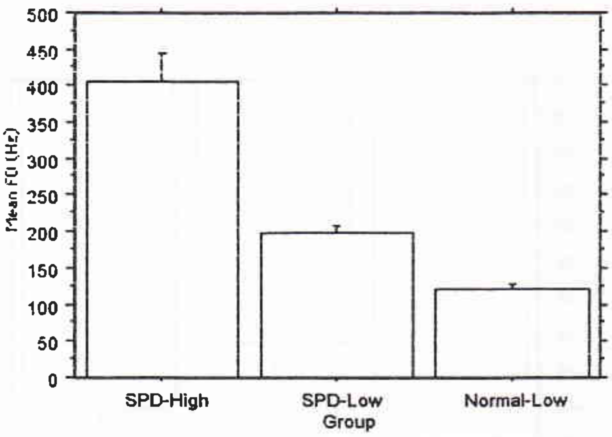

Figure 2. The interaction bar plot of mean fundamental frequency (mean Fo - in Hz) of the groups.

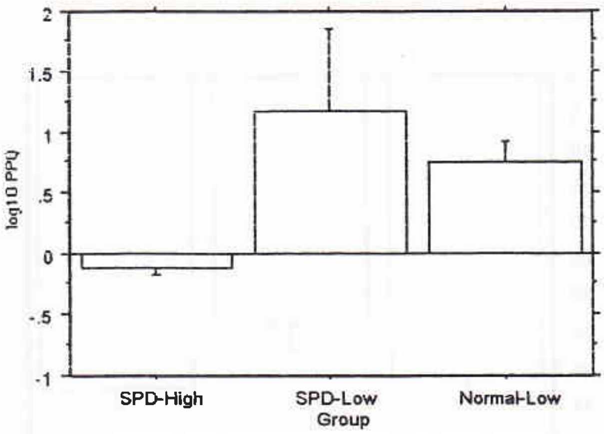

Figure 3. The interaction bar plot of Log10 Pitch Perturbation Quotient (PPQ - %) of the groups.

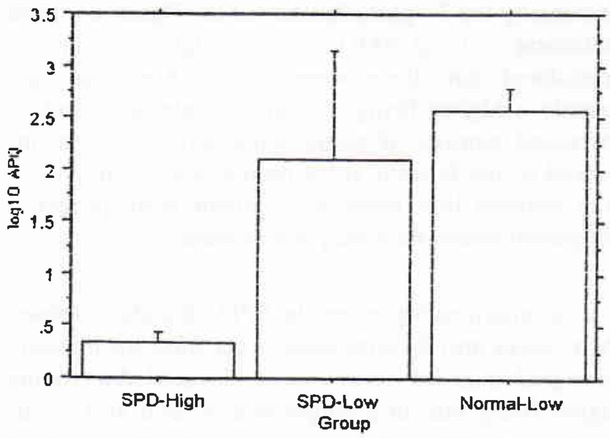

Figure 4. The interaction bar plot of Log10 Amplitude Perturbation Quotient (APQ - %) of the groups.

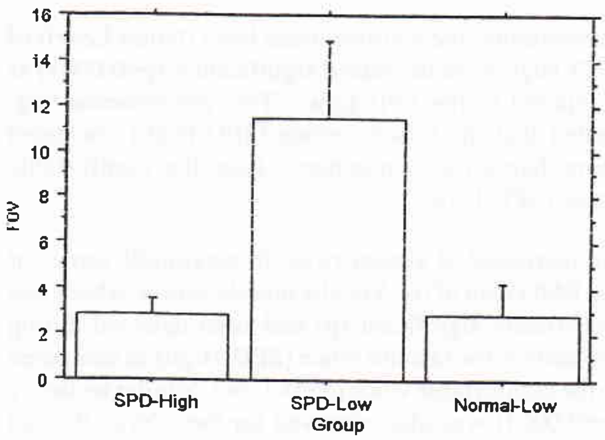

Figure 5. The interaction bar plot of the Fo overall variability (FoV) of the group.

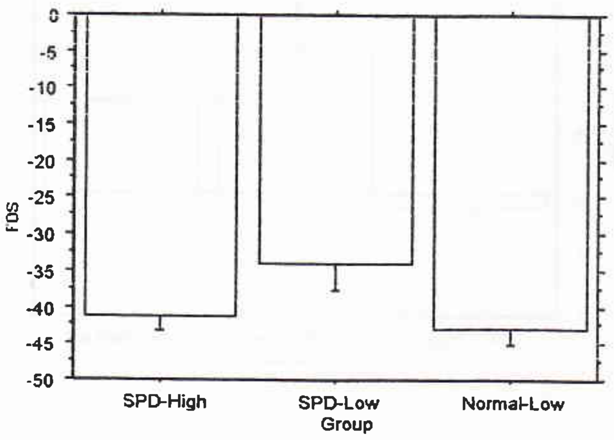

Figure 6. The interaction bar plot of the Slow Fo Perturbations of the group.

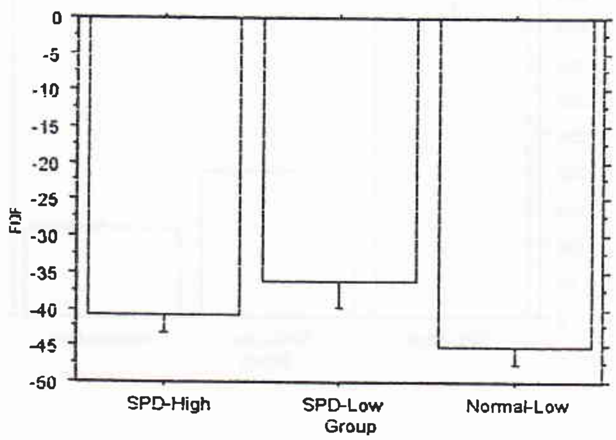

Figure 7. The interaction bar plot of the Fast Fo Perturbations of the group.

pared to the SPD-Low. However, those values were larger than the scores obtained in normal voices, which suggested that SPD voices were less stable and more fluctuating than normal ones. Thus, it was found that the SPD-High showing a better vocal controllability.

Furthermore, the additive noise level (Noise Level) of SPD-High were decreased significantly ($p<0.0001$) as compared to the SPD-Low. This phenomenon suggested that the falsetto voice (SPD-High) contained more harmonic components than the comfortable voice (SPD-Low).

An increased of almost twice in magnitude score of the EMG data of the Vocalis muscle action, which was statistically significant (p) had been detected during production the falsetto voice (SPD-High) as compared to the comfortable voice (SPD-Low). Similar tendency ($p=0.0001$) was also observed for the Cryco-thyroid muscle activation (Figures 10 and 11).

Based on the neuro-physiological aspects, the were 3 mechanisms related to the elevation of muscle contraction activity. First, it is due to the increasing number of its muscle units which are activated. Secondly, by the higher frequency of the impulses in each motor unit, and the last by the synchronization of different motor units.[11] While, Titze[12] with his model experiment stated that the increases in number of motor units and the mean motor unit firing rate, have an effect in decreasing the Fo perturbation value. Based on those statement and our EMG and quantitative datas, we speculated that there were at least 2 mechanisms, namely a higher firing rate in a single unit, and an increased number of motor units activated were involved in the falsetto voice production. Moreover, it was seemed that these mechanism were purely a peripheral neuro-motoric phenomenon.

It is summarized that when the SPD subjects increased their voices into the stipulated target-tone, the intrinsic laryngeal muscles action were elevated due to the higher firing rate in a single muscle unit and an increased number of motor unit activation, which resulted in a better condition of vocal controllability of their voices. It was shown that the acoustic analysis system used in this study was an effective method in evaluating the vocal controllability in either pathologic and normal voice conditions.

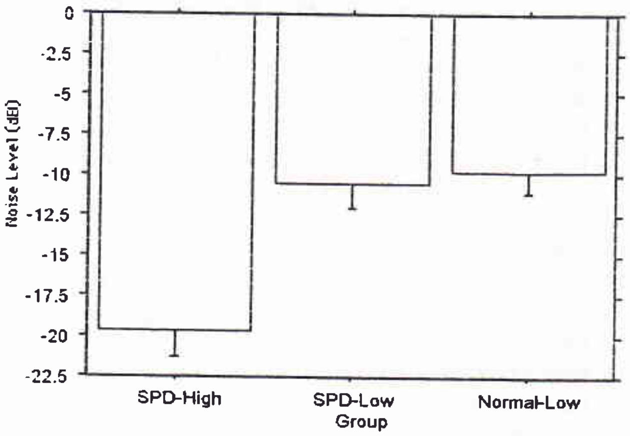

Figure 8. The interaction bar plot of the additive noise level (Noise Level in dB) of the group..

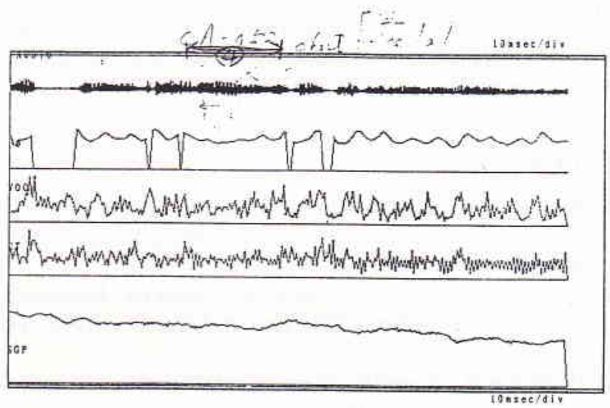

Figure 9. Recording of EMG for Vocalis and Crycothyroid muscles, altogether with voice recording of acoustic wave-form and the F_0 time series. The second half of chosen segment was also shown.

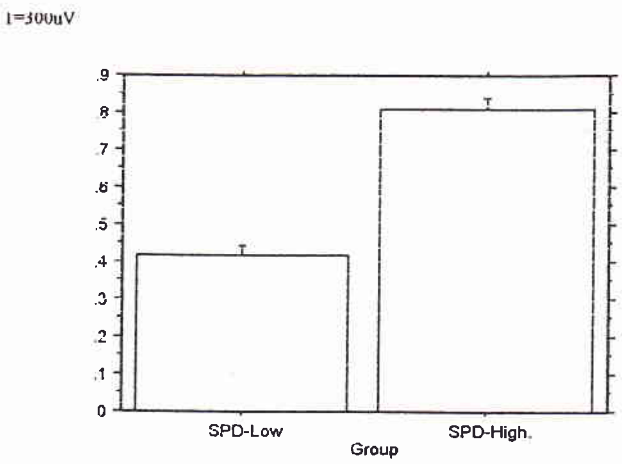

Figure 10. The interaction bar plot of the EMG qualitative data of Vocalis muscle activation on SPD-High and SPD-Low voices.

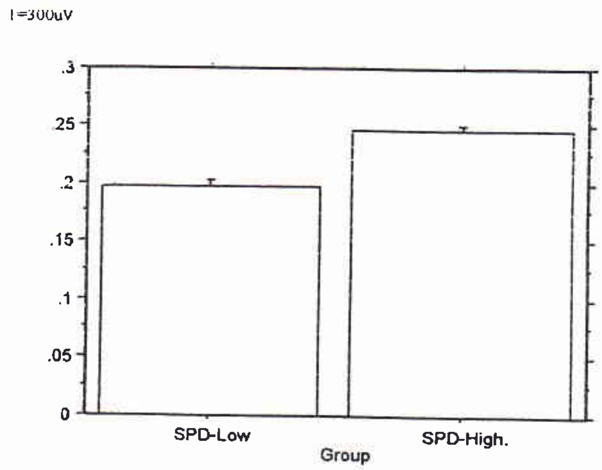

Figure 11. The interaction bar plot of the EMG quantitative data of Cryco-thyroid muscle activation on SPD-High and SPD-Low voices.

The information mentioned above would appear to have a wide range of applicability, particularly in different pitch ranges or phonation modes. However, it is not yet clear whether these results which were observed in sustained vowel phonation were inherent in running speech voice samples. Current research being undertaken in order to demonstrates the acoustical analysis of running speech voice with different phonation modes in SPD subjects.

REFERENCES

1. Abdoerrachman H, Imaizumi S, Hirose H, Niimi S. Slow and Fast Perturbations in Voice. - A Preliminary Report -. Ann Bull RILP 1993 ; 27 : 125-34.
2. Ludlow CL, Bassich CJ, Connor NP, Coulter DC. Phonatory Characteristics of Vocal Fold Tremor. J Phonetics 1988 ; 14 : 509-15.
3. Imaizumi S. Acoustic measures of Pathological Voice Quality. J Phonetics 1985 ; 21 : 457-62.
4. Imaizumi S. Acoustic measurements of Pathological Voice Qualities for Medical Purposes. Proc. ICASSP 1986 ; 1 : 677-80.
5. Abdoerrachman H, Imaizumi S, Niimi S. Test of Vocal Controllability with Increasing Pitch and Intensity. Ann Bull RILP 1994 ; 28 : 45-50.
6. Hirose H, Imaizumi S. Voice Quality in Patients with Neurological Disorders. Vocal Fold Physiology : Voice Quality Control 1994 ; 1 : 235-48.
7. Imaizumi S, Gaufin J. Acoustical Perceptual Characteristics of Pathological Voices : rough, creak, fry and diplophonia. Ann Bull RILP 1991 ; 25 : 109-19.
8. Koike Y. Vowel Amplitude Modulations in Patients with Laryngeal diseases. J Acoust Soc Amer 1969 ; 45 : 839-44.
9. Lim J, Oppenheimer AV, Baida LV. Evaluation of an Adaptive Comb Filtering method for Enhanching Speech Degraded by White Noise Addition. IEEE Trans ASSP 1978 ; 26 : 354-58.
10. Hirose H, Gay T. The Activity of the Intrinsic Laryngeal muscles in Voicing Control : Electromyographic study. Phonetica 1972 ; 25 : 140-64.
11. Hirano M. Clinical Examination of Voice Disorders. Human Communication. 1981 ; 5 : 11-24.
12. Titze IR. A Model for Neurologic Sources of A Periodicity in Vocal Fold Vibration. J Speech Hear Res 1991 ; 34 : 460-72.