

후두질환에 대한 수술 전·후 음성 특성의 음향적 특성 비교 분석*

Analysis and Comparisons of Acoustical Characteristics of Pathologic Voice before and after Surgery

김대현** · 조철우*** · 백무진**** · 왕수건*****

(Dae-Hyun Kim · Cheol-Woo Jo · Moo-Jin Baek · Soo-Geun Wang)

ABSTRACT

In this paper the acoustic characteristics of pathological voice, which are measured before and after surgical operation, are compared. This experiment is conducted for the purpose of predicting patients' speech after operation. The voices are recorded from the same patients. Jitter, shimmer and other parameters are computed and their statistical characteristics are compared. Also spectral changes, such as formant frequency shift and spectral slope change, are compared. From the experimental results, it is verified that not only source characteristics but also vocal tract components vary. And this indicates that the modification of source parameters are not enough for the prediction. Also the result indicates that the operation causes change to both the physical shape of vocal folds and the manner of articulation.

Keywords : Pathologic voice, Acoustic characteristics, Formant frequency, Jitter, Shimmer

1. 서 론

최근 들어 인간의 건강에 대한 관심이 점점 증가하고 있다. 음성을 통한 의사 전달이 사회생활에서의 그 중요성은 기능의 소실 혹은 장애가 왔을 때 절실히 느끼게 된다. 또한 음성은 개인마다 생김새의 차이만큼, 특징적인 음성을 가지고 있어 음성만으로도 개개인을 구별할 수 있는 특이한 부분이다. 따라서 후두질환으로 인한 음성의 장애 혹은 상실은 환자로 하

* 이 논문은 창원대학교 교내연구비 지원과제 '음성신호처리에 의한 양성후두질환자의 수술 후 음성예측에 관한 연구'의 일부로 수행되었습니다.

** 창원대학교 제어계측공학과 박사과정

*** 창원대학교 제어계측공학과 교수

**** 인제대학교 부산백병원 이비인후과 전임강사

***** 부산대학교 의과대학 이비인후과 부교수

여금 의사 소통 및 사회생활에 막대한 지장을 초래하게 되어 삶의 질에 현저한 저하를 가져온다. 그래서 병적 음성으로 치료를 받는 환자의 최대 관심사는 병변의 완전한 제거와 함께 수술 후 음성이 과연 정상적인 음성으로 돌아올 수 있는지, 그렇지 않다면 어느 정도까지 예측이 가능한지에 대해서 많은 의문을 가지게 된다. 특히 생명과 직접 연관이 없는 후두양성 질환에 의한 음성장애가 비수술적 요법으로 실패한 경우는 더욱 그러하다. 수술을 받는 모든 환자들은 음성장애 전에 자신이 가졌던 정상 상태로의 회복을 원할 것이다. 실제 결과는 이와 차이가 날 수도 있고 혹은 거의 정상상태로 회복할 수도 있지만 이를 예측하는 것은 이제까지의 경험과 문헌에 기술된 정보를 이용하여 환자에게 제시할 수밖에 없었다.

어떤 후두 질환의 경우 특징적인 음성을 나타내기 때문에 잘 훈련된 전문가는 환자의 음성만 청취하고도 후두의 상태를 예측할 수 있다는 보고도 있어 음성에 의한 병변의 조기 진단과 감별 진단의 가능성을 보여 다수의 연구자들에 의해 후두질환을 가진 환자의 음성에 대한 연구가 활발히 진행되었다.[1] 그러나 병적 음성에 대한 초기의 연구들은 주로 음성에 대한 청각학적인 방법을 주로 이용하여 병적음성을 분류하고 이를 음향학적 특징 및 특정 후두 질환과 연관지어 객관적인 분류를 하고자 많은 연구가 시행되었다.[2][3][4] 그러나 이는 다소 주관적인 경향이 있었고, 검사자의 용어 및 개념에 따라 진단에 차이를 보였다. 음향분석에 의한 음성 검사는 환자의 음성을 분석하여 후두질환의 진단에 사용하는 것으로 음성이 후두의 여러 가지 생리적, 병태적인 양상에 크게 관련하고 있다는 사실을 근거로 음성 검사 분석 장비의 변수들을 통하여 음성을 분석을 하는 것으로, 현재 임상에서 많이 사용되고 있으며 분석결과를 객관적으로 정량화하기 위하여 각 개인별, 각 질환별에 따른 비교를 하려는 연구가 계속되고 있고 지금까지 어느 정도 공동화된 변수로써 음성검사분석이 가능해지고 있다.

본 논문에서는 먼저 후두질환들의 음성 특성에 대해 알아본 후, 술후 음성의 예측을 위해 필요한 음성학적 지표의 변화를 측정하였다. 병적 음성에 대한 술전 음성과 술후 음성을 음향분석기기를 이용하여 정량적으로 분석하여 술전과 술후의 음성을 비교하고 술후의 음성 개선과 관련된 변수의 변화를 구하였다. 아울러 술전과 술후 음성사이에서 나타나는 포먼트의 변화를 이용하여 음성의 특성에 어떤 변화가 있는지도 비교 분석하고자 한다.

2. 장애음성의 음향적인 특성

음성에 장애가 일어나는 원인의 대부분은 후두의 잘못된 사용과 관련이 있으며, 특히 기능적 음성장애의 경우 계속적인 헛기침, 소리지르기, 울기, 잦은 목 가다듬기 등과 같은 성대 남용이나 부적절한 강도나 음도로 발성하는 성대의 오용으로부터 기인된다. 이러한 성대 오용 및 남용과 함께 위생적인 사용이 이루어지지 않을 경우 성대결절이나 폴립과 같은 기질적인 문제로까지 악화될 수 있다. 일반적인 음성장애환자들에서 나타나는 장애로는 크게 강도장애, 음도장애, 음질장애로 구분할 수 있다. 강도장애는 보통 목소리가 너무 약해서 잘 들리지 않거나 혹은 부적절하게 큰소리가 나는 것을 말한다. 음도장애는 성대결절, 폴립, 성대비대와 같은 성대의 부가적인 조직 때문이다. 마지막으로 음질장애는 발성시 성대를 최적의 상태로 접촉하지 못하기 때문이다. 음질장애는 기식화된 소리, 거친 소리, 목원 소리와 같이

환자의 음성이 정상음성과 구별될 만큼 일탈되는 경우를 말한다. 이러한 음질장애는 양쪽 성대가 견고하면서도 부드럽게 접촉할 때 개선이 가능하다. 위와 같이 장애음성환자들의 음성의 특성은 피치의 주기성이 없어지며, 고주파 성분 즉 잡음성분이 증가한다.[5]

3. 장애음성의 분석 및 통계처리 방법

3.1 장애 음성 수집

장애음성은 1998년 1월부터 2000년 3월까지 음성장애를 주로 부산대학교병원 이비인후과에 내원하여 음성의 개선을 위해 후두미세수술을 시행 받은 환자 중 술후 조직 검사상 양성후두질환으로 확진이 되었으며, 술전 및 술후에 음성검사가 가능하였던 환자를 대상으로 30명 정도를 대상으로 하였다. 음성수집은 방음실에서 편안한 자세로 앉게 한 수 마이크 앞에서 15 cm 가량 거리를 두도록 하여 가능하면 평상시와 같은 높이와 크기로 약 3초간 /아/, /에/, /이/, /오/, /우/를 지속적으로 발성하도록 하였으며, 이를 3회 이상 실시하여 Sony사의 DAT(DTC-59ESJ)에 녹음하였다. 녹음된 음성을 다시 웨이브 파일로 저장하여 분석 가능하게 하였다.

3.2 장애 음성의 성도특성 분석

술전 및 술후 음성을 청취하여 비교했을 경우, 술전 음성은 잡음이 포함된 것 같은 음질 장애가 있었다. 그러나 술후 음성은 정상음성과 비슷하거나 혹은 약간의 잡음성분이 섞인 음성으로 서로 다른 사람으로 판단되는 경우가 많았다. 그래서 음원만을 분석하기보다는 성도특성을 분석해야 할 것으로 판단되었다.

장애음성의 성도특성 분석을 위해서는 선형예측 분석방법을 사용하였다. 이 방법을 통해서 스펙트럴 포락과 제1, 제2, 제3 포먼트 성분을 구해서 비교하였다. 그림 1은 선형예측 스펙트럼과 FFT스펙트럼을 비교한 것이다.

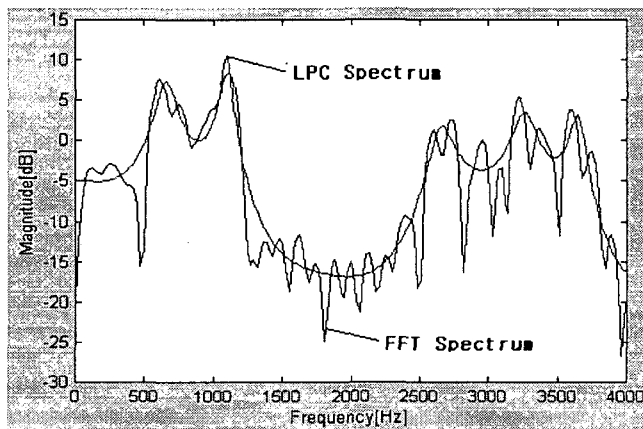


그림 1. LPC 스펙트럼과 FFT 스펙트럼 비교

3.3. 장애 음성파라미터 특성 파라미터 통계처리

녹음된 음성을 Kay사(Kay Elemetrics Corp.)의 Computerized Speech Lab 4300 B(이하 CSL)의 음성 분석 소프트웨어인 Muti-Dimensional Voice Program(MDVP)을 이용하여 총 33 개의 변수에 대하여 분석하여, 장단기 주파수 동요에 관한 변수(Jita, Jitt, RAP, PPQ, sPPQ, vF0), 장단기 진폭의 동요에 관한 변수(ShdB, Shim, APQ, sAPQ, vAm), 소음과 관련된 변수(NHR, VTI, SPI)로 14 개의 변수에 대해 통계처리를 통해 비교 분석했다.[6] 그 중 대표적인 파라미터인 Jitter(Jitt)와 Shimmer(Shim), NHR에 대해 알아보면, Jitter는 피치 주기의 변화율을 나타내는데 사용되어지는 파라미터로 연속적인 피치주기사이의 평균 퍼센트 변화로 나타낼 수 있다. Shimmer는 jitter와는 달리 peak 사이의 크기에 대한 변화율을 나타낸다. NHR은 일반적으로 분석된 신호에서의 소음유무를 나타내주며, 70-4,500 Hz 사이의 배음성분의 에너지와 1,500-4,500 Hz 범위의 배음 이외의 성분에너지의 평균 비율로 나타낸다.

4. 실험 및 결과

먼저 장애음성 파라미터 14 개를 수집된 음성 30명분의 5 개의 모음(/a/, /i/, /u/, /e/, /o/)에 대해 각 파라미터들의 평균값들과 함께 각 파라미터 기준치와의 비율 및 수술후의 개선 정도를 알아보기 위해 술전 술후음성의 파라미터값의 차이로 식 4-1을 이용하여 측정했다.

$$\text{개선율}[\%] = \frac{\text{술전값} - \text{술후값}}{\text{술전값}} \times 100 \quad (4-1)$$

측정값은 특징에 따라 표 1, 2, 3으로 나누어 표시하였다.

표 1. 주파수 동요에 관한 변수

구 분		모 음					전체 평균값	기준치와 의 비율
		/아/	/에/	/이/	/오/	/우/		
Jita	술전	230.30	271.31	216.48	170.45	203.31	218.38	2.62
	술후	88.58	151.91	118.46	74.92	114.19	109.81	1.32
	개선율	61.54	43.64	45.28	56.04	43.83	49.71	
Jitt	술전	3.084	3.648	3.070	2.350	2.990	3.027	2.91
	술후	1.152	1.829	1.725	0.844	1.370	1.382	1.33
	개선율	62.65	49.85	44.14	64.09	54.09	54.35	
RAP	술전	1.748	2.109	1.686	1.270	1.668	1.696	2.49
	술후	0.591	1.009	0.885	0.390	0.767	0.728	1.07
	개선율	66.20	52.12	47.48	69.29	54.02	57.04	

구 분		모음					전체 평균값	기준치와 의 비율
		/아/	/에/	/이/	/오/	/우/		
PPQ	술전	1.817	2.083	1.593	1.243	1.597	1.667	1.98
	술후	0.596	0.966	0.944	0.418	0.611	0.707	0.84
	개선율	67.22	53.62	40.72	66.38	61.77	57.58	
sPPQ	술전	1.488	1.658	2.112	1.432	3.963	2.131	2.09
	술후	1.141	1.299	0.758	0.597	0.783	0.915	0.90
	개선율	23.37	21.62	64.12	58.33	80.24	57.03	
vF0	술전	4.145	4.340	3.889	3.915	6.277	4.514	4.10
	술후	2.481	2.755	3.230	2.500	2.413	2.674	2.43
	개선율	40.15	36.57	16.94	36.35	61.55	40.76	

표 2. 진폭의 동요에 관한 변수

구 분		모음					전체 평균값	기준치와 의 비율
		/아/	/에/	/이/	/오/	/우/		
ShdB	술전	0.949	0.737	0.659	0.554	1.677	0.916	2.61
	술후	0.512	0.566	0.435	0.334	0.329	0.437	1.25
	개선율	45.44	23.14	33.97	39.66	80.36	52.29	
Shim	술전	8.205	8.162	6.908	6.060	7.786	7.424	1.95
	술후	4.421	4.603	3.749	3.115	3.077	3.793	0.99
	개선율	46.12	43.61	45.73	48.60	60.48	48.91	
APQ	술전	5.836	5.386	5.049	4.484	4.574	5.066	1.65
	술후	3.447	4.276	2.999	2.620	2.126	3.094	1.00
	개선율	40.93	20.61	40.60	41.58	53.51	38.93	
sAPQ	술전	5.778	6.003	5.818	5.569	5.990	5.831	1.38
	술후	6.667	4.749	2.991	3.082	2.993	4.096	0.97
	개선율	-15.40	20.61	40.60	41.58	53.51	38.93	
vAm	술전	13.899	12.678	12.126	12.050	12.478	12.640	1.54
	술후	10.758	10.822	10.150	9.441	7.760	9.787	1.19
	개선율	22.59	14.64	16.31	21.65	37.76	22.61	

표 3. 진폭의 동요에 관한 변수

구 분		모음					전체 평균값	기준치와 의 비율
		/아/	/에/	/이/	/오/	/우/		
NHR	술전	0.173	0.183	0.147	0.152	0.183	0.168	0.18
	술후	0.157	0.143	0.128	0.137	0.122	0.138	0.15
	개선율	9.33	21.79	12.36	10.50	33.53	18.07	
VTI	술전	0.090	0.114	0.090	0.043	0.037	0.074	1.22
	술후	0.061	0.074	0.070	0.034	0.026	0.053	0.86
	개선율	33.09	35.59	19.66	21.42	30.79	29.14	
SPI	술전	14.012	10.042	16.392	40.140	60.948	28.307	2.00
	술후	12.601	9.441	4.547	37.425	4.168	27.636	1.96
	개선율	10.072	5.988	11.254	6.762	-5.282	2.3686	

음성분석기기를 이용하여 술전 음성을 분석한 결과 각 변수별 정상 기준치를 기준으로 술전의 비정상 정도를 기준치에 대한 술전 음성의 변수값의 비로 결과를 비교하였다. 술후 음성의 개선에 따라 음성의 잡음성분이 많이 제거되어 음성분석기기에 의한 정량적 분석 결과 각 변수의 수치가 전반적으로 감소하는(정상에 가까워짐) 경향을 보였다.

선형예측 스펙트럼을 이용한 포먼트 값들의 변화는 수집된 음성을 먼저 컴퓨터에서 분석 가능한 툴(Praat 및 Matlab)을 이용하기 위해 웨이브 파일로 변환한 후 결과를 얻었다. DAT에서 받아들일 때 샘플링주파수를 48,000 Hz로 하였으며, 양자화 비트수는 16 비트로 하였다.

각 모음에 대한 술전과 술후 음성에 대한 포먼트 값의 차이(술전음성 포먼트-술후음성포먼트)의 평균값을 표 4에 나타내었으며, 술전 음성의 포먼트가 술후 음성에서 어떻게 변하는지를 알기 위해 부호를 사용 (+)인 경우는 술전이 큰 값, (-)인 경우는 술후가 큰 포먼트 값을 가진다. 제 1, 2포먼트는 대체로 술전음성이 크며, 제 3포먼트의 경우에는 술후음성이 전부 크게 나왔다. 그러나 표준편차의 절대치가 매우 크다는 것에서 개인에 따라 변화의 경향이 일정하지 않다는 것을 알 수 있다. 또 표준편차의 분포에 의해 포먼트 값의 변화 경향은 고차 포먼트의 값의 변화가 더 크다는 것을 알 수 있다.

표 4. LPC 스펙트럼을 이용한 포먼트 값의 차이

구 분		/아/	/에/	/이/	/오/	/우/
F1	AVG	-28.8	-21.3	2.5	7.7	30.9
	STD	137.1	80.8	45.9	64.8	66.6
F2	AVG	24.8	49.8	-34.6	21	-15.1
	STD	153.1	84.3	158.9	106.1	84.5
F3	AVG	-5.8	-39.3	-76.1	-38.2	-71.8
	STD	205.1	168.8	175.46	176.7	212.6

표 4에서는 평균을 했기 때문에 술전 값에서 술후 값으로의 변화가 크지 않기 때문에 그림 2에 평균과 함께 각각의 변화량을 나타내기 위해 표준편차도 함께 나타내었다. 그림 2에서는 제 1포먼트보다 제 2, 3포먼트 각각의 변화량이 크다는 것을 알 수 있다.

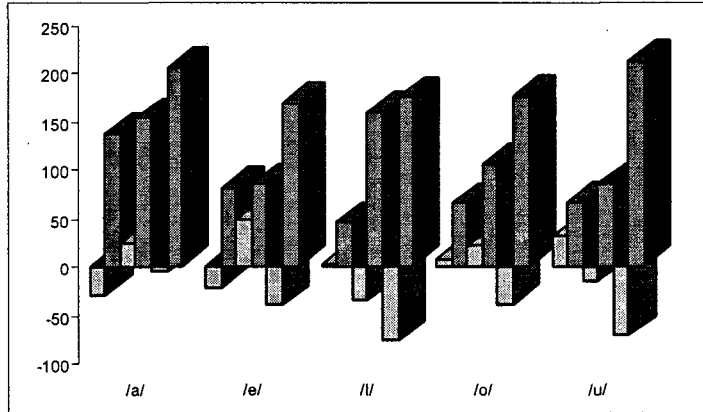


그림 2. 술전 술후 포먼트 차이의 평균과 표준편차

(홀수번: 평균, 짝수번: 표준편차)

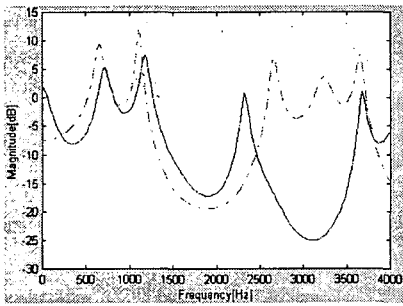
[차이 = 술전포먼트 - 술후포먼트]

개별화자의 특성변화를 관찰하기 위해 두 명의 화자를 선택하여 개별 포먼트 값의 변화와 스펙트럴 포락의 변화를 다섯 개의 모음에 대하여 표 5와 그림 3, 그림 4에 나타내었다. 앞서의 전체적인 통계에서 일치되는 경향은 발견하기 힘들었으나 개별화자의 스펙트럼 변화를 별도로 관찰한 결과 고차 포먼트 값의 상승이 관찰되었고 고대역 스펙트럼의 하강현상이 관찰되었다. 다른 화자에 있어서도 정도의 편차는 크지만 동일한 현상이 발견되었다. 이러한 현상은 화자의 성도특성의 변화를 나타내 주는 것으로 발성패턴이 변했음을 보여준다.

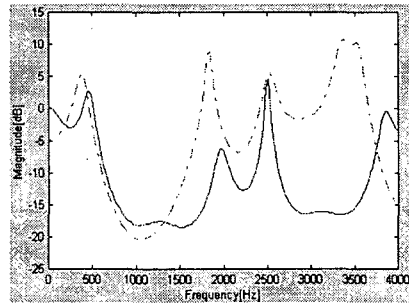
표 5. 화자별 포먼트 값의 변화 및 차이

화자 1	/아/		/에/		/이/		/오/		/우/	
	술전	술후	술전	술후	술전	술후	술전	술후	술전	술후
F1	679	565	492	398	304	281	468	468	351	304
	114		94		23		0		47	
F2	1,171	1,101	1,968	1,828	2,226	2,062	820	773	750	726
	70		140		164		47		24	
F3	2,343	2,671	2,484	2,531	2,906	2,859	2,460	2,625	2,531	2,414
	-328		-47		47		-165		117	

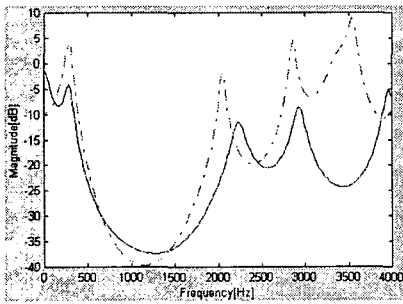
화자 2	/아/		/에/		/이/		/오/		/우/	
	술전	술후	술전	술후	술전	술후	술전	술후	술전	술후
F1	726	750	328	447	234	257	388	388	187	281
	-24		-119		-23		0		-94	
F2	1,223	1,283	1,701	1,641	2,059	2,179	925	865	1,148	867
	-60		60		-120		60		281	
F3	2,417	2,776	2,477	2,835	2,835	3,074	2,358	2,656	2,343	2,367
	-359		-358		-239		-298		-24	



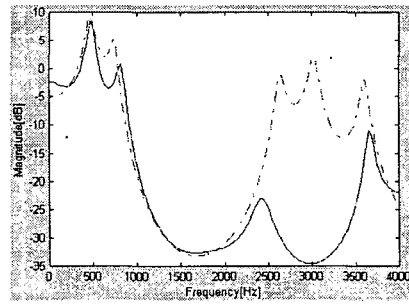
/아/



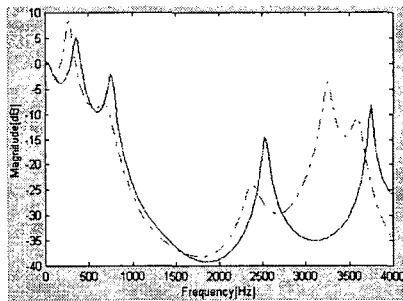
/에/



/이/



/오/



/우/

그림 3. 화자 1의 LPC 스펙트럼 (실선: 술전, 점선: 술후)

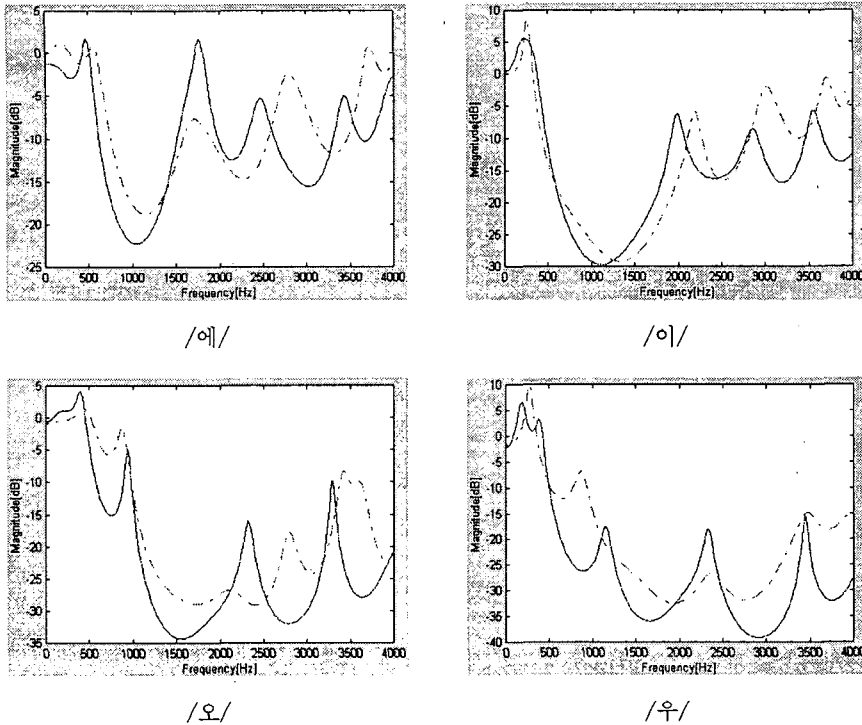


그림 4. 화자 2의 LPC 스펙트럼 (실선: 술전, 점선: 술후)

5. 결론

본 논문에서는 병원에서 내원한 후두질환 환자의 수술 전의 음성과 수술 후의 음성을 수집한 후, 그들 음성간의 특성을 알아보기 위해 음성분석기기인 MDVP를 이용하여 음성특성 파라미터로 14 개 변수에 대해 통계처리를 하였으며, 또한 선형예측 스펙트럼을 이용해 포먼트의 변화도 측정하여 비교 분석했다.

주파수와 진폭 및 잡음에 대한 파라미터들 모두 술후 음성이 많은 개선을 가져왔으며, Jitter와 shimmer, 그리고 NHR이 다른 파라미터에 비해 술전 및 술후 음성의 특성을 대표하는 지표로 측정되었다.

또한 포먼트의 변화도 측정되어, 후두질환 환자 즉 성대질환에 따른 수술에 의해 음성신호의 음원뿐만 아니라 성도 부분도 변화가 온다는 점을 관찰할 수 있었으며 이것은 단순히 음원, 필터모델에서 음원만 변경시켜 주어서는 술후 음성을 예측할 수 없다는 점을 나타내 주고 있다. 이러한 결과는 수술로 인하여 물리적인 구조만 변화하는 것이 아니라 환자의 발성방식 등에도 변화가 있음을 알 수 있었다. 또한 개별 포먼트 값의 변화에 있어서는 환자별 편차가 크게 나타나 일정한 경향이 없음을 알 수 있다. 다만 고차 포먼트 값의 변화가 더 심하게 변한다는 사실을 알 수 있었다.

향후 술전 음성을 변환하여 술후 음성에 가까운 음성으로 재현을 위해 이러한 특성을 고려하여 합성을 위한 정량적 분석을 시도할 것이다.

참 고 문 헌

- [1] Koike.Y, Takhashi.H, Calcaterra.TC, "Acoustic measurements for detecting laryngeal pathology", *Acta Otolaryngol*, 85, 105-17, 1977.
- [2] Issiki.N, Okamura.H, Tanabe.M, Morimoto.M, "Differential diagnosis of hoarseness", *Folia Phoniat*. 21, 9-19, 1969.
- [3] Murry.T, Singh.S, Sargent.M, "Multidimensional classification of abnormal voice qualities", *Jour. Acoust. Soc. Am*, 61, 1630-1635, 1977.
- [4] Hammarberg.B, Fritzell.B, Gaffin.J, Sundberg.J, Wedin.L, "Perceptual and acoustic correlates of abnormal voice quality", *Acta Otolaryngol*, 90, 441-451, 1967.
- [5] 조철우, 김대현, "음원분석을 통한 장애음성의 음향적 특성분석에 관한 연구", *한국음향학회 하계학술발표대회 논문집*, 163-166, 1998.
- [6] *Operations Manual*, "Multi-Dimensional Voice Program(MDVP)", Model 4305, Kay Elemetrics Corp, 1993

접수일자: 2000. 7. 25.

게재결정: 2000. 8. 30.

▲ 김 대 현

경남 창원시 사림동 9번지
창원대학교 제어계측공학과 박사과정(우 : 641-773)
Tel : +82-55-279-7559
e-mail : midas03@taegu.net

▲ 조 철 우

경남 창원시 사림동 9번지
창원대학교 제어계측공학과 교수(우 : 641-773)
Tel : +82-55-279-7552
e-mail : cwjo@sarim.changwon.ac.kr

▲ 백 무 진

부산시 진구 개금동 633-165
인제대학교 부산백병원 이비인후과 전임강사(우 : 614-735)
Tel : +82-51-240-7331, Fax : +82-51-246-8668
e-mail : mjbaek@ijn.inje.ac.kr

▲ 왕수건

부산광역시 서구 아미동 1-10
부산대학교 의과대학 이비인후과 부교수(우 : 602-739)
Tel : +82-51-240-7331, Fax : +82-51-246-8668
e-mail : wangsg@hyowon.cc.pusan.ac.kr