

다채널 음성분석장치를 이용한 정상 성인에서의 발성 방식에 따른 음성변수 분석

서울대학교 의과대학 이비인후과학교실,* 의공학교실**

이상준* · 김광현* · 성명훈* · 노종렬*

권택균* · 이강진* · 박광석** · 최종민**

= Abstract =

Analysis of Voice Parameters on Different Phonatory Tasks using Multi-Channel Phonatory Function Analyzer in Healthy Adults

Sang-Joon Lee, M.D.,* Kwang Hyun Kim, M.D., Ph.D.,* Myung-Whun Sung, M.D., Ph.D.,*

Jong-Lyel Roh, M.D.,* Taek Kyun Kwon, M.D.,* Kang Jin Lee, M.D.,*

Kwang Suk Park, M.S.,** Jong Min Choi, Ph.D.**

Department of Otolaryngology-Head and Neck Surgery, Medical Engineering,**
Seoul National University College of Medicine, Seoul, Korea*

Background and Objectives : The complex physiologic structure of the larynx can vibrate in three or more different ways that yield acoustically and perceptually distinct vocal quality. The purpose of this study is to examine the normal range of voice parameters in Multi-Channel Phonatory Function Analyzer and investigate the difference of voice parameters according to the phonatory patterns.

Materials and Methods : Forty normal adult speakers (20 men and 20 women) with age ranging from third to fourth decades pronounce low, comfortable, and high tone /a/ ; comfortable tone /æ/, /i/, /o/, and /u/ ; fry, falsetto. Voice was analyzed by Newly developed multi-channel phonatory function analyzer.

Results : The normal range of voice parameters in this system was similar to the existing data. Fry shows high jitter and falsetto low SQ. Fry and falsetto show low OQ in men but no difference in women. Jitter, OQ and SQ were different between men and women in modal register, whereas there was no gender difference in fry and falsetto. In frequency magnitude spectrum and EGG, modal register, fry and falsetto have distinguishing pattern.

Conclusions : Modal register, fry and falsetto are distinguishable in voice parameters and show different vibratory patterns.

KEY WORDS : Multi-channel phonatory function analyzer · Fry · Falsetto.

논문접수일 : 2002년 8월 22일

심사완료일 : 2002년 10월 14일

책임저자 : 성명훈, 110-744 서울 종로구 연건동 28번지 5 서울대학교 의과대학 이비인후과학교실

전화 : (02) 760-2447 · 전송 : (02) 745-2387 E-mail : mwsung@snu.ac.kr

서론

성대는 독특한 해부학적, 생리학적으로 인하여 여러 가지 다른 방식으로 진동할 수 있어 음향학적, 지각적으로 다른 음성특질을 만들어 낼 수 있다. 이렇게 서로 다른 음성특질을 Vocal register라고 하며 현재 크게 1) Modal register 2) Falsetto 또는 Loft register 3) Vocal fry 또는 Pulse register 세 가지로 나뉘어진다. Modal register는 최빈값(most common value)이라는 통계적인 용어에서 유래하였다. 보통의 음성과 노래하는 목소리로 chest register, head register가 포함된다. Falsetto는 발성의 고음역에 해당한다. 성대의 긴장도가 최대인 상태로 성대가 매우 얇아져 수직면으로 볼 때 매우 좁은 면에서만 접촉하거나 거의 접촉하지 않는 상태에서 나오는 발성음역이다.¹⁾ Vocal fry는 ‘팝콘튀기는 소리’, ‘모터보트의 엔진소리’로 표현되는 발성음역으로 성대의 긴장도가 거의 없는 상태에서 진동한다.²⁾ 성대의 방사선 영상 검사상 후두실이 매우 감소되어 있거나 거의 존재하지 않는 소견을 보여 가성대가 진성대의 상부에 거의 얽혀진 상태에서 나오는 발성이다.³⁾ Fry는 피치나 강도가 감소되기 시작하는 문장의 끝부분에 흔히 나타난다.²⁾

본 연구의 목적은 성인 남녀에서 다채널 음성분석장치를 이용하여 음성변수의 정상치를 측정하고, 발성 방식에 따라 음성변수에 어떤 차이가 있는지를 알아보고자 하였다.

방법

1. 대상 및 방법

성대에 이상이 없는 것으로 확인된 남녀 각각 20명을 대상으로 하였다. 연령은 20세 이상에서 40세 이하로 한정하였고 평균연령은 30.1세였다. 낮은 소리 /아/, 편안한 소리 /아/, /에/, /이/, /오/, /우/, 높은 소리 /아/, Fry /아/, Falsetto /아/를 발성하게 하였다.

첫째, 편안한 소리 /아/에서 음성 변수들, 즉 기본주파수(fundamental frequency, F0), jitter, shimmer, open quotient(OQ), speed quotient(SQ)의 정상치를 측정하였다. 둘째, 기본주파수에 따라 음성 변수들의 차이가 있는지 알아보기 위해 낮은 소리, 편안한 소리, 높은 소리 /아/를 비교하였다. 셋째, 모음 발성에 따른 음성 변

수들의 차이가 있는지 알아보기 위해 편안한 소리 /아/, /에/, /이/, /오/, /우/간의 비교를 하였다. 넷째, 발성 방식에 따라 음성 변수들의 차이가 있는지 알아보기 위해 편안한 소리 /아/, Fry /아/, Falsetto /아/간의 비교를 하였다. 다섯째, register별로 남녀간에 음성변수들의 차이가 있는지 분석하였다.

2. 음성분석장치

본 연구에서 음성을 획득하고 분석하기 위하여 사용한 다채널 음성분석장치는 음성신호, EGG신호, 유속을 동시에 측정하도록 고안하였다. 음성 신호를 측정할 수 있는 뉴모타코메타(pneumotachometer)가 내장된 측정 기구를 직접 제작하였다. 이 기구를 통하여 음성 신호와 유속을 동시에 측정할 수 있으며, EGG센서도 함께 부착하여 세 개의 신호가 시간적으로 일치하게 된다. 음성을 획득한 후 분석하기 원하는 부분의 구획을 설정하면 분석결과가 나온다(Fig. 1). 채널 1에 음성신호, 채널 2에 EGG 신호, 채널 3에 유속을 보여준다.

3. 음성 변수

음성 신호와 EGG 신호에서 사용되는 변수로는 기본주파수, jitter, shimmer, OQ, SQ 등이 있다. 기본주파수를 구하는 방법에는 크게 주파수 영역에서 구하는 방법과 시간 영역에서 구하는 방법이 있다. 정확한 기본주파수를 구하기 위하여 우리는 두 가지 방법을 모두 사용하였다. 실시간에 가까운 신호처리를 위해서 시간 영역에서 구하는 방법 중에서 Average Magnitude Difference

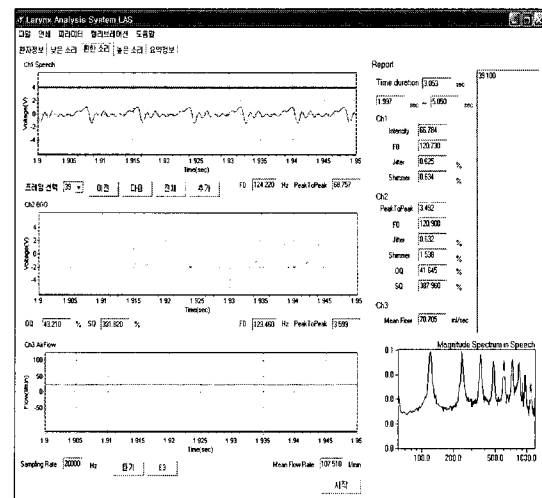


Fig. 1. Interface of multichannel phonatory function analyzer.

Function(AMDF)을 선택했다. AMDF는 자기상관 분석법의 변형된 형태로서 수식으로 표현하면 다음과 같다.

$$AMDF(t) = \frac{1}{L} \sum_{i=1}^L |s(i) - s(i-t)|$$

위 식에서 알수 있듯이 AMDF는 자기상관함수의 곱셈 형태를 뺄셈의 형태로 바꾼 것이라 할 수 있다. 그로 인해 자기상관함수 분석법에 비해 훨씬 빠른 속도를 보여준다. 주파수 영역에서 구하는 방법으로는 Fast Fourier Transform(FFT)를 사용하였다. AMDF는 좋은 해상도를 가지지만, 기본주파수가 두 배로 측정되는 문제가 생길 수 있고, FFT는 50ms 정도의 짧은 시간에서 수행할 경우 해상도가 떨어지는 문제점이 있다. 따라서 FFT로 기본주파수의 범위를 대략적으로 알아내고 AMDF를 이용하여 정확한 기본주파수를 찾아낸다(Fig. 2).

각 프레임에서 구한 기본주파수를 이용하여 jitter를 구하게 되고, 마찬가지로 각 프레임의 진폭을 이용하여 shimmer를 구한다. 정확한 수식은 아래와 같다.

$$Jitter = \frac{\frac{1}{N-1} \sum_{i=1}^{N-1} |T_0(i) - T_0(i+1)|}{\frac{1}{N} \sum_{i=1}^N T_0(i)} \times 100(\%)$$

$$Shimmer = \frac{\frac{1}{N-1} \sum_{i=1}^{N-1} |A(i) - A(i+1)|}{\frac{1}{N} \sum_{i=1}^N A(i)} \times 100(\%)$$

OQ, SQ는 EGG에서 얻어지는 음성 변수이다. OQ는 전체 주기 중에서 열려 있는 기간의 비율로 EGG 그래프상의 10%이상 되는 부분의 시간(tc-t0)을 전체주기(T)로 나눈 값이다. SQ는 성대가 닫히는 시간과 열리는 시간의 비율로 EGG 그래프상 90%에서 10%로 되는 부분의 시간(tc-tb)과 10%에서 90%로 올라가는 부분의 시간(ta-t0)비이다(Fig. 3).

결 과

편안한 소리 /아/에서 측정된 남녀의 음성 변수의 측정

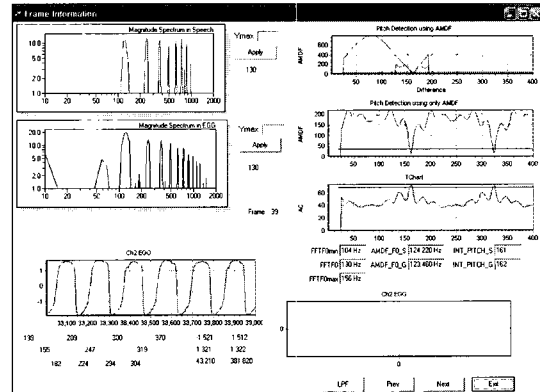


Fig. 2. Accurate pitch detection using both average magnitude difference function (AMDF) and Fast Fourier Transform (FFT).

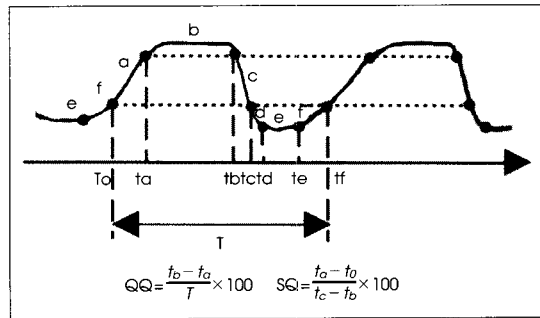


Fig. 3. Open quotient(OQ) and speed quotient(SQ) from electrographic signal. OQ is ratio of the time of open phase to the length of entire vibratory cycle and SQ is ratio between opening time and closing time.

Table 1. Normal range of comfortable tone /a/

	Male	Female
F0(Hz)	130 ±21.5	216 ±16.5
Jitter(%)	0.61 ± 0.15	0.40 ± 0.15
Shimmer(%)	0.61 ± 0.22	0.46 ± 0.34
OQ(%)	34.8 ± 9.29	21.9 ± 6.09
SQ(%)	256 ±97.0	171 ±49.1
Air flow(ml/s)	115 ±60.4	72.0 ±29.0

값은 기존에 알려져 있던 정상치 이내의 범위에 들었다(Table 1).

낮은 소리, 편안한 소리, 높은 소리 /아/ 발성군 사이에 음성 변수의 차이가 있는지 ANOVA(A three-way analysis of variance)로 분석하였다. 남성의 경우 jitter에서 p-value가 0.44로 세 군간에 유의한 차이가 있는 것으로 나오나 사후 검정인 post hoc test에서 낮은 소리 /아/와 높은 소리 /아/사이에만 차이가 있었고 낮은

소리 /아/와 편안한 소리/아/ 사이, 편안한 소리 /아/와 높은 소리 /아/ 사이에는 차이가 없었다. Shimmer, OQ, SQ, 유속에서는 기본주파수의 높낮이에 따라 차이가 없었다. 여성의 경우에도 기본주파수에 따른 음성 변수들의 유의한 차이는 없었다(Table 2).

모음 발성의 차이에 따른 비교에서는 남자의 경우 /아/, /예/, /이/, /오/, /우/발성군 간에 jitter, OQ, SQ, 유속에서 차이가 없었고, shimmer에서는 /이/ 발성군이 /아/, /오/ 발성군 보다 유의하게 낮았지만 다른 군들 간에는 차이가 없었다. 여자에서는 jitter, shimmer, OQ, SQ, 유속에 있어서 모음 발성에 따른 차이는 없었다(Table 3).

발성방식에 따른 음성 변수들의 차이를 알아보기 위해 편안한 소리 /아/, Fry /아/, Falsetto /아/에서 음성

변수를 ANOVA로 분석하였다. Fry의 경우 남녀 모두에서 modal register보다 높은 jitter값을 보였고 falsetto의 경우 남녀 모두에서 낮은 SQ를 보였다. Fry와 falsetto는 높은 shimmer값을 보였으나 통계적으로 유의하지 않았다. Fry는 남성에서 modal register에 비해 낮은 OQ를 보였으며 여성에서도 낮은 OQ를 보였으나 통계적으로 유의하지는 않았다. Falsetto는 남성에서 modal register에 비해서 낮은 OQ를 보였으나 여성에서는 유의한 차이가 없었다(Table 4). 유속은 남녀 모두에서 modal register(남성=115ml/s 여성=72ml/s)보다 fry 발성시(남성=88ml/s,여성=41ml/s)는 낮았고 falsetto 발성시(남성=189ml/s, 여성=153ml/s)는 높았다.

Modal register, fry, falsetto에서 남녀에 따른 차이

Table 2. Comparison of groups with different F0(low, comfortable, and high tone /a/)

	Low	Comfortable	High	p-value
Male				
Jitter(%)	0.71 ± 0.24*	0.61 ± 0.15	0.55 ± 0.19*	0.04
Shimmer(%)	0.66 ± 0.30	0.61 ± 0.22	0.58 ± 0.27	0.58
OQ(%)	33.5 ± 9.72	34.8 ± 9.29	29.2 ± 11.2	0.20
SQ(%)	260 ± 78.3	257 ± 97.0	206 ± 61.1	0.07
Airflow(ml/s)	114 ± 82.1	115 ± 60.4	130 ± 112	0.81
Female				
Jitter(%)	0.58 ± 0.35	0.40 ± 0.15	0.49 ± 0.36	0.18
Shimmer(%)	0.55 ± 0.50	0.46 ± 0.34	0.44 ± 0.25	0.60
OQ(%)	22.2 ± 6.75	21.9 ± 6.09	18.2 ± 6.35	0.10
SQ(%)	179 ± 59.7	171 ± 49.1	161 ± 72.9	0.36
Airflow(ml/s)	74.4 ± 38.4	72.0 ± 29.0	123 ± 33.4	0.83

* : difference in Post Hoc tests

Table 3. Comparison of groups with different vowel(comfortable tone /a/, /æ/, /i/, /o/, /u/)

	/a/	/æ/	/i/	/o/	/u/	p-value
Male						
Jitter(%)	0.61 ± 0.15	0.57 ± 0.19	0.55 ± 0.13	0.53 ± 0.15	0.51 ± 0.20	0.39
Shimmer(%)	0.61 ± 0.22*	0.44 ± 0.12	0.34 ± 0.13*	0.56 ± 0.34*	0.47 ± 0.19	0.00
OQ(%)	34.8 ± 9.29	34.3 ± 9.51	34.8 ± 9.42	33.6 ± 9.78	35.2 ± 8.85	0.99
SQ(%)	257 ± 97.0	253 ± 85.1	270 ± 75.6	245 ± 79.5	257 ± 75.6	0.92
Airflow(ml/s)	115 ± 60.4	114 ± 69.8	111 ± 69.8	132 ± 65.3	145 ± 70.0	0.46
Female						
Jitter(%)	0.40 ± 0.15	0.36 ± 0.18	0.45 ± 0.20	0.38 ± 0.16	0.47 ± 0.32	0.44
Shimmer(%)	0.46 ± 0.34	0.34 ± 0.15	0.25 ± 0.11	0.41 ± 0.16	0.35 ± 0.24	0.04
OQ(%)	21.9 ± 6.09	23.7 ± 7.78	25.1 ± 7.52	26.0 ± 16.4	25.1 ± 12.2	0.77
SQ(%)	171 ± 49.1	161 ± 72.9	165 ± 61.1	158 ± 59.5	164 ± 51.7	0.97
Airflow(ml/s)	72.0 ± 29.0	67.9 ± 47.3	64.1 ± 40.9	69.5 ± 35.4	93.5 ± 61.8	0.39

* : /i/-/a/, /i/-/o/ difference in Post Hoc tests

Table 4. Comparison of groups with different register (comfortable tone /a/, fry, falsetto)

	Comfortable	Fry	Falsetto	p-value
Male				
Jitter (%)	0.61 ± 0.15	1.55 ± 1.22*	0.61 ± 0.24	0.00
Shimmer (%)	0.61 ± 0.22	0.90 ± 0.50	0.78 ± 0.33	0.05
OQ (%)	34.8 ± 9.29	22.0 ± 10.0*	21.1 ± 9.21*	0.00
SQ (%)	257 ± 97.0	274 ± 137	144 ± 102*	0.00
Airflow (ml/s)	115 ± 60.4	88.5 ± 86.4*	189 ± 137*	0.01
Female				
Jitter (%)	0.40 ± 0.15	1.37 ± 1.15*	0.60 ± 0.48	0.00
Shimmer (%)	0.46 ± 0.34	0.61 ± 0.27	0.48 ± 0.28	0.29
OQ (%)	21.9 ± 6.09	18.0 ± 8.95	24.1 ± 11.5	0.14
SQ (%)	171 ± 49.1	150 ± 69.9	120 ± 67.8*	0.04
Airflow (ml/s)	72.0 ± 29.0	41.1 ± 36.0	153 ± 137*	0.04

* : difference in Post Hoc test

Table 5. Comparison of voice parameters between male and female in modal register, fry and falsetto

		Modal register		Fry		Falsetto	
			p-value		p-value		p-value
Jitter (%)	M	0.61 ± 0.15	0.00	1.55 ± 1.22	0.65	0.61 ± 0.24	0.96
	F	0.40 ± 0.15		1.37 ± 1.15		0.60 ± 0.48	
Shimmer (%)	M	0.61 ± 0.22	0.11	0.90 ± 0.50	0.05	0.78 ± 0.33	0.00
	F	0.46 ± 0.34		0.61 ± 0.27		0.48 ± 0.28	
OQ (%)	M	34.8 ± 9.29	0.00	22.0 ± 10.0	0.22	21.1 ± 9.21	0.39
	F	21.9 ± 6.09		18.0 ± 8.95		24.1 ± 11.5	
SQ (%)	M	257 ± 97.0	0.00	274 ± 137	0.00	144 ± 102	0.41
	F	171 ± 49.1		150 ± 69.9		120 ± 67.8	

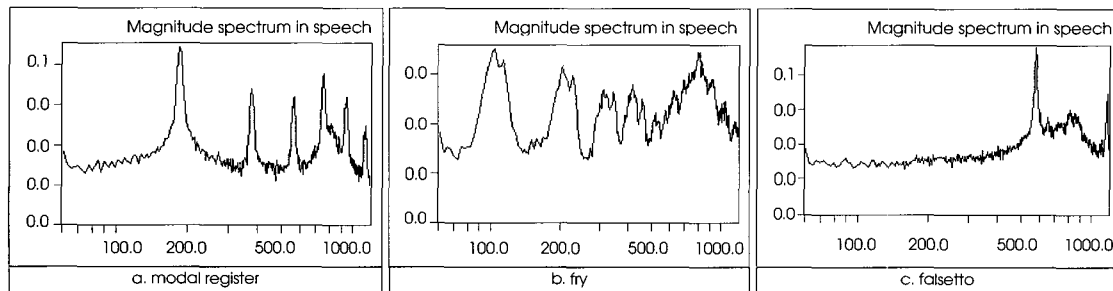


Fig. 3. Open quotient(OQ) and speed quotient(SQ) from electrographic signal. OQ is ratio of the time of open phase to the length of entire vibratory cycle and SQ is ratio between opening time and closing time.

를 알아보기 위해 T-test를 하였다. Modal register의 경우 jitter, OQ, SQ에서 남녀간에 유의한 차이를 보였다. Fry의 경우 jitter가 modal register보다 유의하게 높았지만 남녀간의 비교에서는 차이를 보이지 않았다. 또한 OQ가 fry에서 modal register보다 유의하게 낮았지만 남녀간의 차이는 없었다. Falsetto의 경우 SQ가

modal register보다 유의하게 높았지만 남녀간의 차이는 없었다. 또한 falsetto에서 jitter, shimmer의 남녀간의 차이도 없었다(Table 5).

Falsetto는 EGG 파형상 작은 톱니모양의 파형 이외에도 더 큰 주기로 느린 파형이 보였다(Fig. 5).

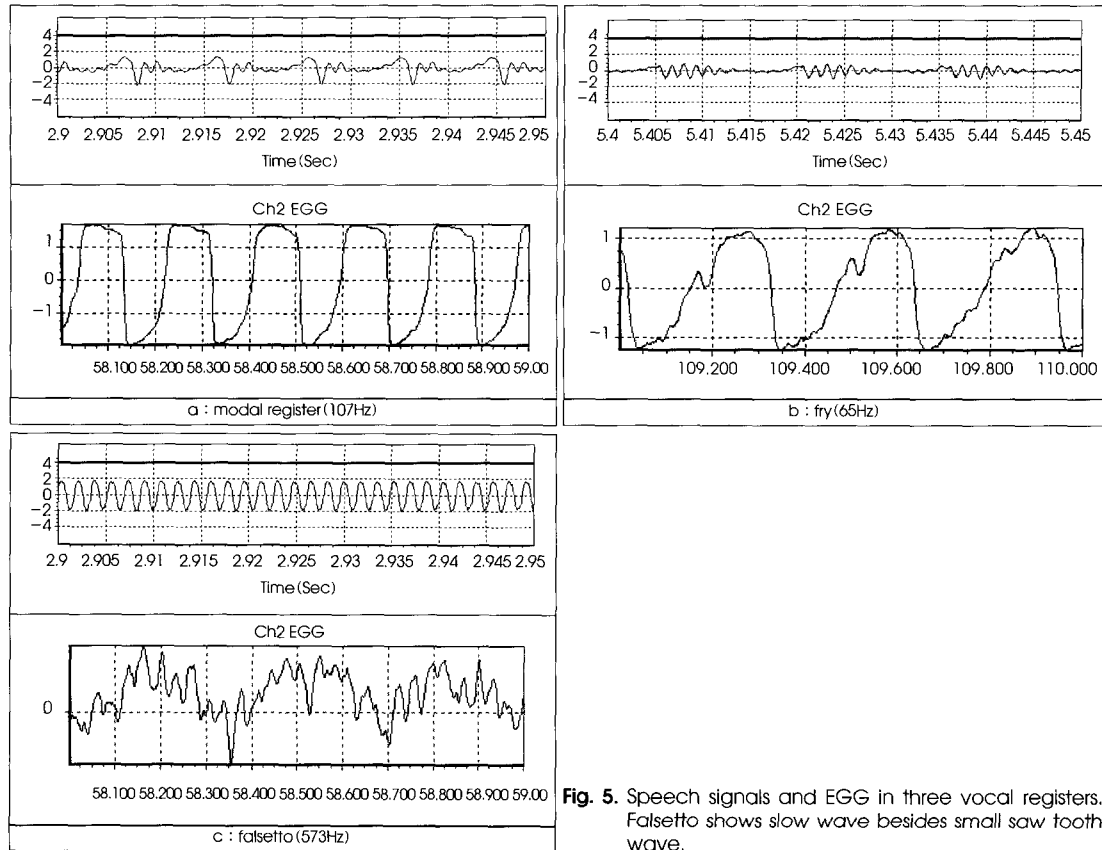


Fig. 5. Speech signals and EGG in three vocal registers. Falsetto shows slow wave besides small saw tooth wave.

고찰

낮은 소리, 편안한 소리, 높은 소리 /아/ 사이, 모음 /아/, /에/, /이/, /오/, /우/사이에 음성변수가 차이가 있는지 ANOVA 분석한 결과 기본주파수의 차이에 따른 jitter, shimmer, OQ, SQ의 차이가 없었고 모음 발성에 따른 차이도 없어 음의 고저나 모음 발음이 달라져도 jitter, shimmer, OQ, SQ는 달라지지 않음을 알 수 있었다.

Fry의 경우 jitter가 유의하게 상승되어 있는데 이는 fry시 발성의 안정성이 감소됨을 의미하며 주파수 스펙트럼 상에도 곡선의 폭이 넓어진 소견을 보였다(Fig. 4). OQ는 fry시 성대의 폐쇄기가 길어지기 때문에 감소된다.²⁾ 가성대는 본질적으로 비근육적인 용적이며 fry시는 진성대와 연속적이어서 발성시 수직방향의 점막파동이 이동할 때 기본적으로 가성대와 진성대와 관련되어 진행하므로 폐쇄기가 길어진다. 여성의 경우는 남성의 경우에 비해 OQ의 감소가 유의하지 않았는데 이는 남

성과 여성이 동일한 기본주파수 범위에서 fry 발성을 하지만, 여성의 경우 남성보다 성대의 두께가 더 얇아도 가능하기 때문으로 생각된다.³⁾ Falsetto시 SQ가 chest register나 fry보다 유의하게 낮았다. Hirano 등⁴⁾은 falsetto발성시 후두내전근의 근전도의 활동도가 낮아진다고 보고하였다. 따라서 후두내전근의 근긴장의 감소로 인하여 성대가 닫히는 시간이 길어지기 때문이다.⁵⁾ Modal register발성시 유속과 기본주파수와의 관계는 논란이 많다. Large 등⁶⁾은 고음시 후두의 내경이 좁아지고 성문이 더 세게 닫힘으로써 유속이 감소한다고 하였고 Rubin 등⁷⁾은 기본주파수가 증가하면 유속이 증가한다고 하였으며 McGlone⁸⁾은 기본주파수와 유속 간에는 별 상관관계가 없었다고 하였다. Fry 발성시 유속이 감소되는데 이것은 성문의 폐쇄기가 증가되는 것과 관련이 있다고 생각되며 falsetto시는 modal register에 비해 유속이 크게 증가된다.²⁾⁸⁾ Fry에서 modal register에 비해 jitter가 높고, OQ가 낮지만 남녀간에 차이가 없었으며 falsetto에서도 SQ가 modal register에 비해 유의하

게 낮지만 남녀간에 차이가 없었다. 이 결과는 modal register와 비교해서 fry와 falsetto는 각각이 남녀모두에서 비슷한 성대의 발성양상을 갖는다는 것을 뒷받침한다.

다채널 음성분석장치는 음성을 각 프레임 단위로 계산하는데 한 프레임의 시간은 0.05초이다. Fig. 5는 modal register, fry, falsetto에서 한 프레임의 음성신호와 EGG신호이며 한 프레임을 각각 10등분해서 보여주고 있다. 음성신호는 초 단위로 표시되어 있고 EGG신호는 프레임으로 표시되어 있다(modal register의 경우 음성신호는 2.9~2.95초까지 이며 EGG신호는 이에 해당되는 58프레임이다). 음성신호, EGG신호 모두에서 modal register는 한 프레임에 5개약간 이상(107Hz), fry는 3개 약간 이상(65Hz), falsetto 28~29개(573Hz)의 파형이 보인다. falsetto에서는 위의 파형이외에 프레임당 3개 정도의 느린 파형이 보이는데 이는 다른 피험자들의 경우에도 비슷한 모양이 나타났다. 발성시 후두의 상하운동에 의한 저항도 증가, 인두내 압력상승에 의한 저항도 증가등 EGG파형에 영향을 줄 수 있는 요인이 많아 정확한 평가는 어렵지만 falsetto시 성대의 직접적인 접촉이 거의 없는 빠른 진동과 함께 더 느린 주기로 성대가 진동하는 것으로 생각된다.⁹⁾

결 론

다채널 음성분석장치를 이용함으로써 음성신호, EGG신호, 유속을 동시에 신속하게 측정할 수 있었다.

Modal register, fry, falsetto는 음성변수값에 있어서도 서로 구별되는 특징이 있었다. Jitter와 OQ가 fry를 modal register와 가장 잘 구별시켜 주는 음성변수였고

SQ가 falsetto를 가장 잘 나타내는 음성변수였다.

중심 단어 : 다채널 음성분석장치 · Fry · Falsetto.

References

- 1) Baken RJ : *An overview of laryngeal function for voice production. In : Sataloff RT : Professional voice. The science and art of clinical care. 2nd ed. San Diego : Singular Publishing Group, Inc., 1997 : 147-165*
- 2) Michael B, Yang C, Manwa LN, Harvey RG : *Acoustic, aerodynamic, physiologic, and perceptual properties of modal and vocal fry registers. J Acoust Soc Am. 1998 ; 103 (5) : 2649-2657*
- 3) Allen EL, Hollien H : *A laminographic study of pulse (vocal fry) register phonation. Folia phoniat. 1973 ; 25 : 241-250*
- 4) Hirano M, Vennard W, Ohala J : *Regulation of register, pitch, and intensity of voice. Folia phoniat. 1970 ; 22 : 1-20*
- 5) Kitzing P : *Photo-and electroglottographic recording of the laryngeal vibratory pattern during different registers. Folia phoniat. 1982 ; 34 : 234-241*
- 6) Large J, Iwata S, Leden H : *The operatic head register versus falsetto. Folia phoniat. 1972 ; 24 : 19-29*
- 7) Rubin H, LeCover M, Vennard W : *Vocal intensity, subglottic pressure and air flow relationship in singers. Folia phoniat. 1967 ; 19 : 393-413*
- 8) McGlone RE : *Air flow in upper register. Folia phoniat. 1970 ; 22 : 231-238*
- 9) 안종복 · 정옥란 : EGG사용 및 분석방법. In : 고도홍 · 정옥란 · 신호근 등 : 음성 및 언어 분석기기 활용법. 한국문화사, 2001 : 255-271