

# A Comparison among Repeated Measures of the Voicing Efficiency Protocol in the Aerodynamic Assessment

Seung Jin Lee<sup>a</sup>, Sung-Eun Lim<sup>b</sup>, Jae-Yol Lim<sup>a</sup>, Hong-Shik Choi<sup>a</sup>

<sup>a</sup>Department of Otorhinolaryngology and the Institute of Logopedics and Phoniatrics, Yonsei University College of Medicine, Seoul, Korea

<sup>b</sup>Voice Clinic, Department of Otorhinolaryngology, Gangnam Severance Hospital, Seoul, Korea

**Correspondence:** Hong-Shik Choi, MD, PhD  
Department of Otorhinolaryngology and the  
Institute of Logopedics and Phoniatrics, Yonsei  
University College of Medicine, 211 Eonju-ro,  
Gangnam-gu, Seoul 06273, Korea  
Tel: +82-2-2019-3461  
Fax: +82-2-3463-4750  
E-mail: hschoi@yuhs.ac

Received: July 6, 2018  
Revised: August 2, 2018  
Accepted: August 2, 2018

**Objectives:** The purpose of the present study was to determine whether voicing efficiency measures derived from the Phonatory Aerodynamic System differ according to repeated trials and the presence of voice disorders upon controlling for severity. In addition, the diagnostic ability of the first and averaged measures were compared. **Methods:** A total of 160 participants (80 patients and 80 controls; 23 males and 57 females for each group) participated in the study. Each participant was asked to perform the voicing efficiency protocol three times. Selected aerodynamic parameters including mean sound pressure level (PHODB), mean pitch (MEAP), mean peak air pressure (MPAP), mean airflow during voicing (MFPHO) were measured. A two-way repeated measures of ANCOVA upon controlling for severity and ROC curve analysis were performed. **Results:** Results indicated that there was no difference in MPAP and MFPHO among repeated trials upon controlling for severity, although a significant difference was found in PHODB (trial 1 < trial 2 = trial 3). The patient group showed lower MFPHO compared to the control group. ROC curve analysis revealed that the AUCs were .683 and .687 for the first and averaged MPAP; and .644 and .625 for the first and averaged MFPHO, respectively, suggesting no difference between measurement methods. **Conclusion:** In sum, the current data showed that MPAP and MFPHO measures did not differ according to repeated trials. Furthermore, the diagnostic ability of the first and averaged measures was similar, providing a rationale for the use of a time-efficient single trial protocol in clinics and research.

**Keywords:** Phonatory aerodynamic system, Voicing efficiency, Repeated measures, Subglottal pressure, Mean airflow rate

음성장애의 평가에는 숙련된 전문가의 청지각적 평가를 비롯하여 음향학적 평가, 공기역학적 평가 등을 이용한 다면적 평가가 권고되어 오고 있다(Ma & Yiu, 2011). 이 중 공기역학적 평가는 호흡유량계(pneumotachograph)가 장착된 기기를 이용하여 측정된 다양한 공기역학적 측정치들을 이용하여 음성 산출 시 가장 중요한 에너지, 즉 호흡의 효율적인 사용 여부를 가늠할 수 있는 평가양식이다. 현재 임상 및 연구에서 공기역학적 평가에 널리 이용되는 기기로는 Phonatory Aerodynamic System Model 6600 (PAS; Kay-Pentax, Montvale, NJ, USA)이 있다. PAS는 대상자의 공기역학적

특성에 대해 조망할 수 있는, 사용자의 편의성을 고려한 다양한 검사 프로토콜들을 제공하고 있으며, 이러한 프로토콜로는 폐활량(vital capacity), 공기압 스크리닝(air pressure screening), 최대연장발성(maximum sustained phonation), 편안한 연장발성(maximum sustained phonation), 음압수준 변화(variation in sound pressure level), 음성 효율성(voicing efficiency), 연결발화(running speech) 등이 있다. 아울러 전기성문파형검사(electroglottography)를 위한 동 제조사의 하드웨어 제품인 KayPENTAX Electroglottograph (EGG) Model 6103을 PAS 외장모듈에 연결하면 음성 산출 시의 성문의 접

측 정도에 대한 데이터를 동시에 기록할 수 있는 프로토콜이 별도로 마련되어 있다는 장점도 있다(KayPENTAX, 2010).

PAS를 이용한 공기역학적 평가를 통해 발성 시 성대의 기능을 나타내는 다양한 용량(volume), 흐름(flow), 압력(pressure), 발성의 효율성 등을 측정할 수 있다(Joshi & Watts, 2016). 주로 이용되는 변수들로는 성문하압(subglottal pressure, cmH<sub>2</sub>O)과 평균호기류율(mean airflow rate, L/s)이 있으며, 이들은 모두 PAS의 음성 효율성 프로토콜(voicing efficiency protocol, VOFE)을 이용하여 측정이 가능하다. VOFE 프로토콜의 결과로 측정되는 변수들 중에서 성문하압은 평균성문하압(mean peak air pressure, MPAP)이라는 변수로, 평균 호기류율은 유성음 산출 시 평균호기류율(mean airflow during voicing, MFPHO)이라는 변수로 산정된다. 그 외에도 유성음 산출 시 평균음압(mean sound pressure level during voicing, PHODB)과 평균음도(mean pitch, MEAP)를 측정할 수 있어 임상적 활용도가 높다고 할 수 있다. 음향학적 측정치에 해당되는 PHODB와 MEAP는 PAS 외장모듈에 외장 마이크를 장착시킴으로써 정확하게 측정이 가능하다. 음도와 음량 또한 공기역학적 특성에 중요한 영향을 미칠 수 있는데, 윤상갑상근(cricothyroid muscle)이 수축하면 성대의 두께를 감소시켜 성문 기류(glottal flow)의 증가, 소음 산출, 노력성 호흡 등을 야기할 수 있다(Zhang, 2017).

그런데 이 VOFE를 시행함에 있어 한번 시행하여 산출된 결과에 대한 해석만으로도 충분하지, 아니면 여러 차례 시행을 반복하여야 하는지에 대해서는 다소 논란의 여지가 있다. 먼저 VOFE를 반복적으로 시행한 연구들을 살펴보면 주로 3회에 걸친 반복 시행을 하며, 이렇게 반복 측정된 수치들의 평균값을 대표적인 값으로 간주한다(Kim, 2014; Ko, Choi, Lim, & Choi, 2015; Schaeffer, 2017; Dargin & Searl, 2015). 이 연구들은 동일 대상자 내에서도 반복 시행에 따라 변이성(variability)이 있을 수 있다는 가정을 전제로 하고 있다. 한국인 정상 성인을 대상으로 한 기준 연구(Kim, 2014)도 이러한 방식을 취하고 있는데, PHODB, MEAP에서 회차 간 유의한 차이가 있었던 반면, MPAP와 MFPHO에서는 회차에 따른 차이는 없으나 안정적인 측정을 위해 3회 반복하여 측정할 것을 권고하였다.

반면에 1회 측정만을 시행한 연구들은 주로 /p<sup>h</sup>a/ 음절을 5-9회 시행토록 하여 이중 맨 앞과 맨 뒤의 음절을 제외한 나머지 가운데의 3-7회의 샘플을 분석하였다(Awan, Novaleski, & Yingling, 2013; Dastolfo, Gartner-Schmidt, Yu, Carnes, & Gillespie, 2016; Kang, Jo, Kim, & Chang, 2017; Liang et al., 2017; Liu, Zheng, Tian, Yang, & Zou, 2015; Zraick, Smith-Olinde, & Shotts, 2012). 이들 연구에서는 대개 음절을 다섯 번 산출하여 가운데의 세 음절을 분석한 3회

반복 프로토콜에 비해 많은 수의 음절을 산출하여 처음과 마지막 음절을 제외하고 나머지 음절들을 분석하는 경향이 관찰된다.

PAS의 공식 매뉴얼을 살펴보면, 자체적인 지침은 VOFE에 대한 설명 부분에 구체적으로 명기되어 있지 않으나, 부록에서 인용한 기준 연구(Zraick et al., 2012)에서는 동일 프로토콜을 3회 실시하여 평균값을 구하였다(KayPENTAX, 2010). 다만 저널에 출판된 Zraick 등(2012)의 절차를 살펴보면 음절을 9회 산출하도록 하여 가운데의 일곱 음절을 분석하였다고 기술한 데 반해, PAS의 매뉴얼에서는 7회 산출하도록 하여 가운데의 다섯 음절을 분석하였다고 기술되어 있어 구체적인 방법론에 있어 다소 혼동을 주고 있다. 또한 현재 임상에서 널리 활용되고 있는 PAS 소프트웨어의 VOFE 프로토콜의 설정상 반복 시행 결과에 대해 자동적으로 평균값을 제시해주는 기능은 구현되어 있지 않으므로, 기준 자료를 임상에서 의미 있게 이용하고자 한다면 수작업으로 계산을 해야만 하는 실정이다.

이와 같은 맥락에서 대부분의 선행연구들에서 공히 나타나는 경향은 1회 시행 혹은 반복 측정에 대해서 배경이나 근거, 혹은 음성장애 환자에서 반복 시행에 따른 임상적 의의가 명확하게 제시되어 있지는 않다는 점을 지적할 수 있다. 이는 아마도 다수의 환자군과 대조군을 이용하여 PAS의 진단적 유용성에 대해 1회 측정과 반복 측정을 이용한 측정치 사이에 비교를 수행한 선행연구의 부재에 기인하는 것으로 여겨진다. 만약 3회에 걸쳐 시행한 값들 간에 서로 유의한 차이가 있거나, 1회치의 측정치와 3회 측정치의 평균값 사이에 진단적 유용성에 있어 차이가 있다면, 반드시 3회 측정을 시행하여 평균값을 구하여 이를 해석하여야 할 것이다. 한국인에 대한 기준 연구(Kim, 2014)에서 제시된 측정치들의 신뢰구간이 모두 이러한 방식으로 산정되었다는 점을 고려하면 더욱 그러하다. 반면에 만약 이러한 차이가 뚜렷하게 관찰되지 않고, 1회 측정치와 반복 측정에 의한 평균치가 갖는 진단적 능력에 있어 큰 차이가 없다면 1회 시행으로 충분하다는 데 대한 실증적이고 과학적인 근거를 제시할 수 있을 것이며, 아울러 시간 대비 효율성이 증시되는 임상현장에서 더욱 효율적인 검사 진행에 도움이 되는 임상적 의의를 가질 수 있을 것으로 여겨진다. 다만 이러한 반복 측정의 의의 혹은 필요성에 대해 엄밀하게 판단하기 위해서는 측정치 자체에 큰 영향을 미칠 수 있는 음성장애 중증도와 같은 요인을 감안하여 연구설계가 이루어져야만 할 것으로 본다.

따라서 본 연구에서는 음성장애 환자군과 정상 대조군에서 VOFE를 3회 반복 실시하여, 청지각적 평가를 통해 측정된 중증도를 통제하였을 때 집단과 회차(trial) 간 PHODB, MEAP, MPAP, MFPHO의 차이가 있는지 알아보고자 하였다. 또한, MPAP와 MFPHO에

서 첫 번째 측정치와 세 측정치의 평균치가 갖는 진단적 유용성에 차이가 있는지 살펴보고자 하였다.

## 연구방법

### 연구대상

먼저 G\*Power (Faul, Erdfelder, Buchner, & Lang, 2009)를 통하여, 집단(2개)과 반복측정(3회)에 따른 비교(two-way repeated measures of ANOVA)를 위한 최소 표본수를 산정한 결과 총 144명이었다(effect size  $f(V) = .3$ , Power = .9). 이에, 환자군으로는 2016년 3월부터 2018년 2월까지 서울의 한 대학병원 이비인후과에 내원하여 이비인후과 전문의에 의해 음성장애로 진단된 만 20세 이상의 환자 80명(남 23명, 여 57명)을 선정하였다. 다만 내원 전 이미 타기관에서 수술적 혹은 행동적 치료를 받은 자는 대상자에서 제외하였다. 환자군의 음성장애 진단명은 편측 성대폴립(vocal polyp) 20명, 성대결절(vocal nodules) 12명, 편측 성대마비(vocal cord palsy) 10명, 내전형 연축성 발성장애(adductor spasmodic dysphonia) 9명, 인후두 역류질환(laryngopharyngeal reflux)과 성대 낭종(vocal cyst) 각 7명, 성대 과각화증(hyperkeratosis of vocal cords)과 성대구증(sulcus vocalis) 각 5명, 라인케씨 부종(Reinke's edema) 2명, 성문 암(glottic cancer), 근긴장성 발성장애(muscle tension dysphonia), 성대 육아종(vocal cord granuloma) 각 1명이었다. 환자의 진단명 가운데 과다기능적(hyperfunctional) 음성장애에 속한 경우는 42명(성대폴립, 성대결절, 내전형 연축성 발성장애, 근긴장성 발성장애), 과소기능적(hypofunctional) 음성장애에 속한 경우는 15명(편측 성대마비, 성대구증)이었다(Korean Society of Laryngology, Phoniatics and Logopedics, 2012).

정상 대조군은 갑상선 수술 전 음성검사가 의뢰되어 내원한 대상자 80명이었다. 대조군의 성별은 환자군과 일치시키되, 연령은 3세 이내로 매칭하여 집단 간 연령 차이는 없었다( $t = -.008, p = .994$ ). 갑상선 수술을 앞둔 대상자에서 잠재적인 음성장애의 가능성을 배제하기 위하여 면담, 연결 발화에 대한 청지각적 평가, 설문지를 통한 심리측정적 평가, 그리고 후두스트로보스코피 검사를 이용하였다(Awan, Roy, & Cohen, 2014; Lee, Choi, & Kim, 2018). 먼저 검사일을 기준 3개월 내에 상기도 감염이나 음성 질환으로 이비인후과를 방문한 이력이 있거나, 과거나 현재의 흡연 이력이 보고된 자, GRBAS 척도(Hirano, 1981)를 이용한 청지각적 평가에서 G 척도상 0이 아닌 자, 음성활동 및 참여 프로파일-한국판(Korean version of the Voice Activity and Participation Profile, K-VAPP)에 응답하도록 하여 총점의 절단점인 14.5점 이상인 대상자는 제외하였

**Table 1.** Demographic data of the participants

Parameter	Patient group (N=80)	Control group (N=80)	p-value
Age (yr)	41.46±9.91 (21-59)	41.45±9.78 (20-59)	.994
Gender			1.000
Male	23 (28.75)	23 (28.75)	
Female	57 (71.25)	57 (71.25)	
Professional voice user	35 (43.75)	45 (56.25)	.101
G	1.95±.57	.00±.00	<.001***
K-VAPP			
Total score (280)	121.23±77.82	.80±2.10	<.001***
ALS (100)	40.41±28.19	.36±1.06	<.001***
PRS (100)	40.24±30.79	.23±1.01	<.001***

Values are presented as mean±SD (range) or number (%). G=grade from the GRBAS scale; K-VAPP=Korean version of the Voice Activity and Participation Profile; ALS=activity limitation score; PRS=participation restriction score. \*\*\* $p < .001$ .

다(Lee et al., 2016; Lee, Pyo, & Choi, 2018). 마지막으로 전문의가 실시한 후두스트로보스코피 검사 결과 후두의 이상 소견을 보이지 않는 경우에만 연구에 포함되었다. 각 집단의 인구학적 특징에 대한 통계치가 Table 1에 제시되어 있다.

### 연구절차

먼저 의무기록 검토와 면담을 통하여 기본정보와 직무상 음성 사용의 정도, 발병시기, 흡연력, 음성장애 치료력 및 관련 병력 등의 기본정보를 수집하였다. 사전 면담 이후 모든 대상자가 수행하는 과제는 크게 심리측정적 평가를 위한 자기 보고식 설문지 과제와 표준 문단 읽기 과제, 공기역학적 평가를 위한 VOEF 과제로 나누어 진행되었다. 먼저 심리측정적 평가로는 K-VAPP (Lee et al., 2016) 설문지에 대상자 본인이 응답하되, 설문지 과제의 목적을 사전에 주지시키고, 11점 척도에 대해 설명하였으며, 본인의 주관적인 느낌을 토대로 하여 검사 당시의 상태를 기준으로 설문을 시행하도록 안내하였다. 이를 통해 K-VAPP의 총점과 활동제한점수(activity limitation score, ALS), 그리고 참여제한점수(participation restriction score, PRS)를 계산하였다. 표준 문단 읽기 과제에서는 '가을' (Kim, 2012) 문단 전체를 읽도록 하되, 대상자로 하여금 자연스러운 음도와 크기로 읽도록 하였으며(Lee, Lim, & Choi, 2017), 이에 대하여 제1저자(1급 언어재활사, 언어병리학 박사, 음성검사, 치료 및 연구 경력 7년)와 2저자(1급 언어재활사, 언어병리학 석사, 음성검사, 치료 및 연구경력 20년)가 GRBAS 척도(Hirano, 1981)로 합치에 이르도록 평정하였다.

공기역학적 평가과제로는 PAS의 VOEF 과제를 시행하였다. 먼



저 PAS의 외장모듈에서 마스크를 분리한 후 1리터 실린더를 결합하여 영점 조정(calibration) 절차를 수행하였고, 직경 1 mm 가량의 일회용 플라스틱 튜브를 마스크에 장착하여 호흡유량계에 연결하였다. 마이크 입력단자에는 AKG CK77 무지향성 콘덴서 마이크 (Harman International, Stamford, CT, USA)를 연결한 후, 고무 와셔를 이용하여 마이크 스탠드에 단단히 고정시켜 기류헤드의 끝 부분으로부터 2 cm 거리를 유지하도록 하였다. 대상자가 의자에 편안하게 착석한 후, PAS 외장모듈의 손잡이를 양손으로 잡고 수평을 유지한 상태에서 마스크를 착용하도록 하였다. 튜브의 끝부분은 혀의 위에 위치하도록 하되, 대상자가 튜브의 끝 부분을 깨물거나, 혀로 막지 않도록 안내하였다. 그 후 마스크의 윗부분을 검사자가 손으로 밀면서 마스크 전체가 대상자의 얼굴에 완전히 밀착되도록 한 상태에서 /p<sup>h</sup>a/를 한 호흡에 가급적 편안한 음도와 크기로 일정하게 5회 반복 산출하도록 하는 과제를 3회 반복 시행하였다 (Ko et al., 2015). 대상자가 사전에 방법을 정확하게 수행할 수 있을 때까지 연습 기회를 제공하였으며, 가급적 한 호흡에 초당 1.5-2회 정도의 속도로 5회를 산출하도록 하되, 5회 산출이 어려울 경우 각 음절 산출 후 짧게 휴기를 하도록 하였다.

### 자료 분석

3회 반복 시행한 각 5개씩의 /p<sup>h</sup>a/ 음절 샘플 중 가운데 세 개씩의 샘플을 선택하여 분석하였다(각각 1, 2, 3회차). 분석에 필요한 구간은 세 음절의 발성 구간 중에서 프로그램에 의해 자동적으로 선택 되도록 하였다. 각 회차의 PHODB (PHODB<sub>1</sub>, PHODB<sub>2</sub>, PHODB<sub>3</sub>), MEAP (MEAP<sub>1</sub>, MEAP<sub>2</sub>, MEAP<sub>3</sub>), MPAP (MPAP<sub>1</sub>, MPAP<sub>2</sub>, MPAP<sub>3</sub>), MFPHO (MFPHO<sub>1</sub>, MFPHO<sub>2</sub>, MFPHO<sub>3</sub>)를 측정하였으며, 변수별로 세 회차의 측정치에 대한 평균값 (PHODB<sub>Avg</sub>, MEAP<sub>Avg</sub>, MPAP<sub>Avg</sub>, MFPHO<sub>Avg</sub>) 또한 계산하였다.

### 통계 분석

중증도(G)를 통제한 상태에서 집단(2개)과 측정 회차(3개)에 따라 각 변수의 차이가 있는지 알아보기 위하여 이원 반복측정 공분산분석(two-way repeated measures of ANCOVA)을 시행하여 윌크의 람다(Wilk's λ)를 산정하였으며, 구형성 가정을 만족하지 못하는 경우에는 그린하우스-가이저(Greenhouse-Geisser)의 자유도 보정 방식을 적용하여 유의수준을 해석하였다. 유의수준은 5%로 하였으며, 사후분석은 본페로니(Bonferroni) 방법을 이용하였다. 첫 회차의 MPAP<sub>1</sub>, MFPHO<sub>1</sub>, 그리고 평균값의 MPAP<sub>Avg</sub>, MFPHO<sub>Avg</sub>가 음성장애 유무에 대해 갖는 진단적 유용성과 측정방법 간 차이를 알아보기 위해 수신자 조작 특성(receiver operating characteristic, ROC) 곡선 분석을 수행하여 곡선 아래 영역(area under curve, AUC)을 계산하여 두 방법 간 비교를 하였다. SPSS 23.0 프로그램 (IBM-SPSS Inc., Armonk, NY, USA)를 이용하여 이원 반복측정 공분산분석을, Med-Calc 통계 소프트웨어 v17.9.7 (MedCalc Software, Ostend, Belgium; <http://www.medcalc.org>)을 이용하여 ROC 곡선 분석을 시행하였다.

### 연구결과

#### 집단 및 회차에 따른 비교

중증도를 통제한 상태에서 집단(2개) 및 반복 회차(3회)에 따른 비교를 수행한 결과, 모든 변수에 대해 교호작용은 없었다(Table 2). MEAP, MPAP의 경우 집단 및 회차에 따른 유의한 차이가 없었던 반면(Figure 1), PHODB의 경우 회차에 따른 차이가 있었다 ( $p = .011$ ). 사후분석 결과 PHODB<sub>1</sub>에 비해 PHODB<sub>2</sub> ( $p = .025$ )와 PHODB<sub>3</sub> ( $p = .003$ )의 측정치가 유의하게 높았고, PHODB<sub>2</sub>와 PHODB<sub>3</sub> 사이에는 유의한 차이가 없었다. 환자군의 MFPHO는 전

**Table 2.** Descriptive data of the acoustic and aerodynamic measures according to groups and trials

Parameter	Group	Trial				Interaction <i>p</i> -value	Within-subject		Between-subject	
		1	2	3	Avg		Wilk's λ	<i>p</i> -value	<i>F</i>	<i>p</i> -value
PHODB (dB)	Patient	80.966±6.444	81.009±6.971	81.281±6.967	81.086±6.635	.635	.957	.011*	.546	.461
	Control	81.240±5.490	82.123±5.499	82.229±5.658	81.865±5.479					
MEAP (Hz)	Patient	170.543±41.519	172.852±44.986	171.075±38.691	171.428±40.317	.050	.966	.066	.116	.734
	Control	166.333±40.544	167.979±41.332	168.677±41.369	167.663±41.030					
MPAP (cmH <sub>2</sub> O)	Patient	7.136±2.125	7.177±2.317	7.169±2.443	7.161±2.174	.576	.991	.401	.979	.324
	Control	5.849±1.781	5.845±1.896	5.761±1.851	5.818±1.761					
MFPHO (L/s)	Patient	.209±.133	.208±.139	.214±.152	.210±.141	.700	.989	.564	6.257	.013*
	Control	.145±.069	.148±.075	.146±.078	.147±.072					

Values are presented as mean ± SD.

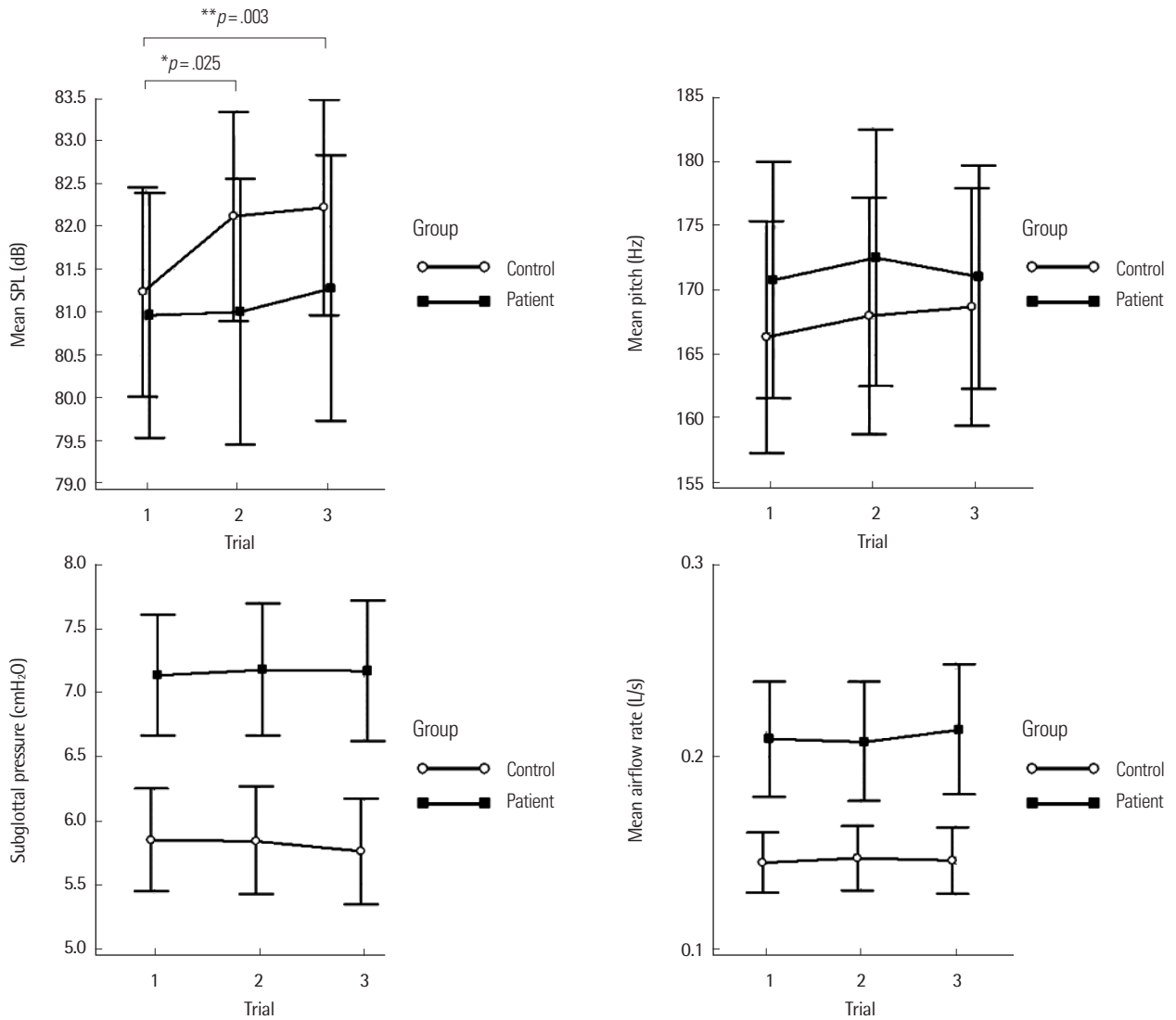
PHODB = mean sound pressure level during voicing; MEAP = mean pitch; MPAP = mean peak air pressure; MFPHO = mean airflow during voicing.

\* $p < .05$ .

반적으로 대조군보다 높았으나, 중증도를 통제한 상태에서는 대조군보다 낮았다( $p = .013$ ).

측정방법에 따른 진단적 유용성

MPAP<sub>1</sub>과 MPAP<sub>Avg</sub>, MFPHO<sub>1</sub>과 MFPHO<sub>Avg</sub>의 진단적 유용성을

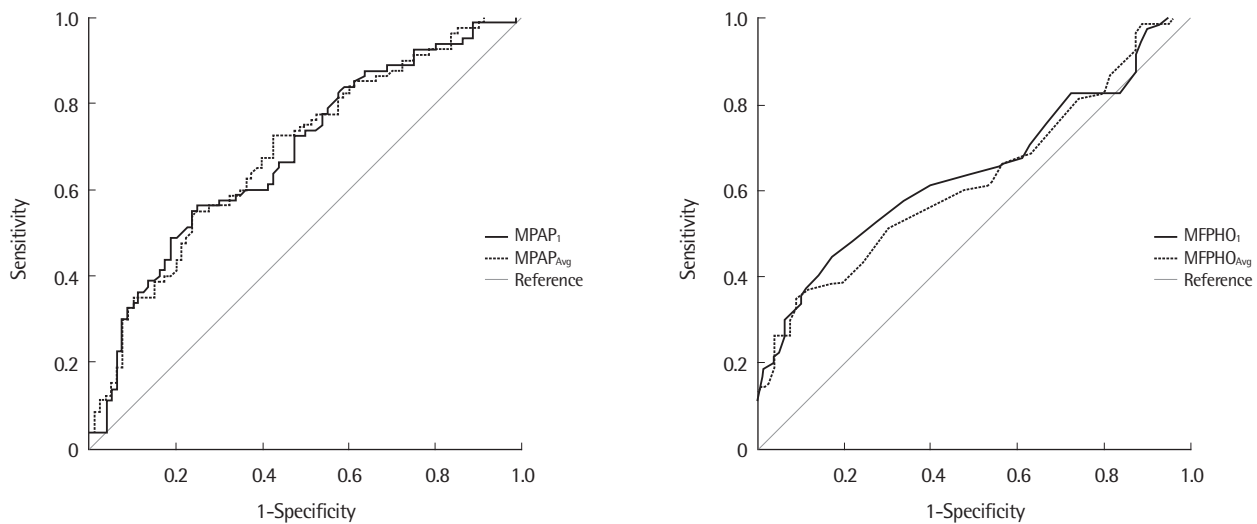


**Figure 1.** Clustered multiple comparison graphs of the acoustic and aerodynamic measures. \* $p < .05$ , \*\* $p < .01$ . SPL = sound pressure level.

**Table 3.** Comparison of the diagnostic ability between the first versus averaged aerodynamic measurement methods

	AUC	95% CI	z	p-value	Cut-off score	Sensitivity	Specificity
MPAP (cmH <sub>2</sub> O)			.226	.821			
MPAP <sub>1</sub>	.683	.601-.765			6.085	.600	.600
MPAP <sub>Avg</sub>	.687	.605-.769			6.120	.625	.625
MFPHO (L/s)			1.320	.187			
MFPHO <sub>1</sub>	.644	.558-.730			.155	.613	.600
MFPHO <sub>Avg</sub>	.625	.538-.711			.165	.550	.625

AUC = area under curve; CI = confidence interval; MPAP = mean peak air pressure; MPAP<sub>1</sub> = first measured MPAP, MPAP<sub>Avg</sub> = average of the MPAP measures among three trials; MFPHO = mean airflow during voicing; MFPHO<sub>1</sub> = first measured MFPHO; MFPHO<sub>Avg</sub> = average of the MFPHO measures among three trials.



**Figure 2.** Receiver operating characteristic (ROC) curve of the aerodynamic measures. MPAP<sub>1</sub> = first measured mean peak air pressure; MPAP<sub>Avg</sub> = average of the mean peak air pressure measures among three trials; MFPHO<sub>1</sub> = first measured mean airflow during voicing; MFPHO<sub>Avg</sub> = average of the mean airflow during voicing measures among three trials.

알아보기 위해 AUC를 산정한 결과, MPAP<sub>1</sub>은 .683 (95% CI, .601-.765), MPAP<sub>Avg</sub>는 .687 (95% CI, .605-.769), MFPHO<sub>1</sub>은 .644 (95% CI, .558-.730), MFPHO<sub>Avg</sub>는 .625 (95% CI, .538-.711)였다(Table 3). 가장 높은 민감도와 특이도를 갖는 절단점은 MPAP<sub>1</sub>은 6.085 cmH<sub>2</sub>O, MPAP<sub>Avg</sub>는 6.120 cmH<sub>2</sub>O, MFPHO<sub>1</sub>은 0.155 L/s, MFPHO<sub>Avg</sub>는 0.165 L/s였다. 첫 번째 측정치와 세 측정치의 평균값 간에 AUC를 비교한 결과, MPAP ( $z = 0.226, p = .821$ )와 MFPHO ( $z = 1.320, p = .187$ ) 모두 유의한 차이가 없었다(Figure 2).

### 논의 및 결론

본 연구에서는 PAS를 활용한 공기역학적 평가, 그 가운데에서도 널리 이용되는 VOEF 프로토콜을 세 차례 실시하여 중증도를 통제 한 상태에서 음향학적, 공기역학적 측정치를 비교하였다. 그 결과 대개의 변수에서 반복 측정 회차에 따른 차이가 없었고, 집단 간 차이도 없었다. 아울러 첫 번째 회차의 측정치와 세 측정치의 평균값의 진단적 유용성도 차이가 없음을 확인하였다.

먼저 세 회차 간의 차이에 대해 살펴보면, 중증도를 통제하였을 때 주요 공기역학적 측정 변수라 할 수 있는 MPAP와 MFPHO에서 회차 간 차이가 없었고, 음향학적 측정치 중 MEAP에서도 통계적으로 유의한 차이는 없었다. 이는 정상 성인에서 반복 회차의 측정치 간 차이가 없었던 국내 선행연구들과 일치하는 결과이다(Kim, 2014; Ko et al., 2015). 또한 이러한 결과는 선행연구 중 1회만 시행하도록 한 연구들을 지지하는 결과이며, 임상에서 시간적으로 효

율성이 높은 1회 프로토콜의 사용을 뒷받침한다(Awan et al., 2013; Dastolfo et al., 2016; Kang et al., 2017; Liang et al., 2017; Liu et al., 2015; Zraick et al., 2012).

다만 PHODB의 경우 첫 번째 회차의 측정치에 비해 두 번째와 세 번째 회차의 측정치가 높았다. 이는 첫 번째 회차 이후 한번 더 산출하라는 지시사항을 들은 대상자가 더욱 좋은 수행력을 보이기 위해 큰 음량으로 과제를 수행하는 경향이 임상 사례들에서 종종 관찰되는 것과 무관하지 않은 것으로 보인다. 국내 선행연구에서도 반복 측정횟수가 증가할수록 PHODB의 측정치가 증가한다고 보고하였으며, 평균 측정치뿐 아니라 산포도, 즉 표준편차 또한 크게 증가하는 것으로 나타났다(Kim, 2014). 특히 첫 번째에서 세 번째 회차로 진행하면서 남성의 경우 3.68 dB에서 8.70 dB로 비교적 크게 증가한 반면, 여성의 경우 4.20 dB에서 4.53 dB로 그 증가폭이 크지 않아 다소의 성별 차가 있을 가능성을 시사하였다. 반복 시행에 따른 측정치 변화의 이러한 성별차에 대해서는 후속연구를 통해 규명할 필요성이 있을 것으로 보인다. 또한 MEAP의 증가 경향은 유의하지는 않았으나 유의수준에 매우 가까웠는데( $p = .050$ ), 이는 3회 반복 측정치 간 비교를 시도한 국내의 선행연구들(Kim, 2014; Ko et al., 2015)에서도 일맥상통하는 결과이며, 음량의 증가 경향과도 무관치 않은 것으로 보인다. 다만 이러한 음량의 증가가 공기역학적 측정치의 차이까지 이어지지 않은 것은 본 연구에서 가급적 편안한 음도와 크기로 과제를 수행하도록 모니터링하였기 때문인 것으로 보인다.

회차별로 MPAP의 평균 측정치들을 살펴보면, 환자군에서 7.136,

7.177, 7.169 cmH<sub>2</sub>O였던 반면, 정상군은 5.849, 5.845, 5.761 cmH<sub>2</sub>O로, 측정치 자체만으로는 전반적으로 환자군에 비해 낮았으나, 중증도를 통제하였을 때에는 집단에 따른 차이가 없었다. 정상군의 이러한 측정치는 Kim (2014)의 국내 기준 연구의 정상 측정치에 비해서도 낮은 편이며, Zraick 등(2012)의 여성 집단의 측정치(평균 5.57 cmH<sub>2</sub>O)와 유사한 수준이다. 이는 한편으로는 본 연구의 대상자 중 여성의 비율이 높았다(71.25%)는 점과 무관하지 않은 것으로 보이며, MPAP에 영향을 줄 수 있는 관련 요인들, 예컨대 무성파열음 조음의 정확도, 음도나 음량의 적절성 여부 등에 대한 후속연구가 정상 대상자뿐만 아니라 음성장애 환자에서도 보다 면밀하게 이루어져야 할 필요가 있다고 본다.

가령 무성 파열음인 /p<sup>h</sup>/를 얼마나 정확하게, 구강내압을 충분히 형성하여 발음하느냐에 따라 MPAP의 측정치에 다소 차이가 있을 수 있다. 모음, 무성파열음, 모음의 연쇄를 발음할 때 성대와 성도(vocal tract)는 공기역학적인 상호작용을 하게 되는데, 파열음 조음을 위해 성도의 끝 부분이 폐쇄되면 성대 진동이 끝나는 성문의 면적이 수동적으로 증가하게 되며(Delebecque, Pelorson, & Beauteemps, 2016), 무성파열음의 정확한 조음이 이루어지지 않으면 이러한 상호작용에 영향을 주어 공기역학적 측정치도 달라질 수 있다는 것이다. 실제로 임상에서는 /p<sup>h</sup>/를 유성음화하거나 약화시켜 발음함으로써 결과치에 영향을 주는 사례가 종종 관찰되며(Choi & Choi, 2014), 우리말의 양순 평음 파열음이 모음 사이의 어중초성 위치에서 유성음화되는 점을 생각해 보면 이러한 가능성은 더욱 커진다. 이와 더불어 환자의 긴장도나 임상가의 구체적인 지시사항의 적절성이 영향을 미쳤을 가능성을 배제할 수 없다. 검사어의 발성 유형에 따른 측정치 차이를 살펴 본 Choi와 Choi (2014)의 연구결과를 보면, MFPHO의 경우 /p<sup>h</sup>/와 /p/ 사이에는 유의한 차이가 없었지만, MPAP의 경우 /p<sup>h</sup>/가 더 높은 측정치를 보였다. 또한 음소 자체에 따른 음도나 음량의 차이가 있을 수 있으므로, 이를 감안한 정확한 검사 적용이 필요할 것으로 보인다.

중증도를 통제한 상태에서 집단 간 공기역학적 측정치의 비교 결과를 보면, MPAP의 경우 차이가 없었던 반면, MFPHO의 경우 환자군이 더 낮은 측정치를 보였다. 이는 공기역학적 평가에 있어서 다양한 측정치들을 종합적으로 살펴보아야 할 필요성과 더불어 음성장애 연구에 있어 중증도에 대한 통제 여부가 매우 중요한 요인이 될 수 있음을 뒷받침하는 결과인 것으로 보인다(Lee et al., 2017). 이는 중증도 자체 만으로도 이미 연구의 대상이 되는 종속변수에 지대한 영향을 끼치기 때문에, 이를 통제하지 않는다면 관심의 대상이 되는 변수가 순수하게 종속변수에 미치는 영향을 살펴보기 어렵기 때문이다. 중증도를 통제하였을 때의 이러한 차이는 절단점

에 대한 해석 또한 주의를 요한다는 점을 시사한다. 단순하게 생각하면 MPAP와 MFPHO의 측정치는 환자군이 대조군에 비해 높았으므로, 절단점 이상의 측정치를 보이면 음성장애 집단으로 선별 가능한 것으로 생각할 수 있다. 그러나 중증도나 혹은 과다기능, 과소기능적 음성장애의 특성을 생각하면 절단점을 사용한 이러한 구분만으로 대상자의 음성장애 여부를 판단하는 것은 다소 성급할 수 있다는 것이다.

MPAP와 MFPHO 모두 진단적 능력에 있어서는 첫 번째 측정치와 세 측정치의 평균값 간에 유의한 차이가 없었다. Kim (2014)의 연구에서는 세 번의 반복 측정치 간의 차이가 없음에도 불구하고, 정상 성인에 대한 연구였다는 점에서 조심스럽게 매뉴얼에 따른 3회 반복 측정을 권고한 것으로 여겨지나, 본 연구의 이러한 결과는 세 측정치 간의 비교 결과와 더불어 1회 측정을 지지하는 것이라 할 수 있다. 다만 구체적으로 AUC 값들을 살펴보면 모두 양호(fair) 수준에 그쳤으며, 이는 1회 측정의 근거가 약하기 때문이라고 하기보다는 임상에서 공기역학적 측정치에 대한 해석 시 주의가 필요하다는 점, 그리고 다양한 검사도구들을 통한 다면적 음성장애 평가와 종합적인 해석이 중요하다는 점을 시사한다고 본다(Ma & Yiu, 2011). 다만 어떠한 음성장애 환자군에서 공기역학적 검사결과를 어떠한 식으로 해석할지, 그 임상적 시사점이 무엇인지에 대해서는 개별 질환군에 따른 후속연구가 필요할 것이다. 예컨대 다른 여러 검사양식들을 통해 불완전한 성대 접촉이 확인된 경우, 일반적으로 MFPHO가 증가해야 함에도 불구하고 공기역학적 검사를 통해 본 결과 MFPHO가 특별히 크게 증가하지 않았다면, 환자의 보상적인 노력(compensatory efforts)과 그에 기인한 이차적인 근긴장성 발성장애(muscle tension dysphonia) 등의 존재 가능성을 시사하는 것이라고 볼 수도 있을 것이다. 아울러 개별 변수들의 측정치뿐만 아니라 이들을 조합하여 과다 혹은 과소기능적 음성장애 등에 대한 새로운 지표를 만들 수 있는 가능성 또한 후속연구에서 탐색해 볼 직하다.

요약하자면 MPAP와 MFPHO는 음성장애 여부와 반복 시행에 따른 차이를 보이지 않았으며, 진단적 능력 또한 1회 측정값과 3회 평균값의 차이를 보이지 않았으므로, 임상 및 연구에서 VOEF 프로토콜을 시행할 때 시간상 보다 효율적인 1회 검사로 충분할 것이라는 결론을 내릴 수 있다.

다만 본 연구에서는 과다기능 혹은 과소기능적 음성장애 간에 발생할 수 있는 차이를 고려하여 세분화된 비교를 수행하지는 못하였으며, 성별 간의 차이나 연령층에 따른 차이를 연구에 포함하지 못하였다는 한계가 매우 명백하다. 따라서 추후 서로 다른 다양한 병인의 음성장애 환자에 대해 성별과 노화에 따른 차이를 고려



한 연구가 이루어져야 그 임상적 유용성을 더할 수 있을 것으로 판단된다. 아울러 공기역학적 검사결과가 구체적으로 어떠한 검사양식과 결합하였을 때 유용한 임상적 정보를 제공할 수 있는지에 대한 심층적인 후속연구가 필요할 것으로 여겨진다.

## REFERENCES

- Awan, S. N., Novaleski, C. K., & Yingling, J. R. (2013). Test-retest reliability for aerodynamic measures of voice. *Journal of Voice*, 27, 674-684.
- Awan, S. N., Roy, N., & Cohen, S. M. (2014). Exploring the relationship between spectral and cepstral measures of voice and the Voice Handicap Index (VHI). *Journal of Voice*, 28, 430-439.
- Choi, S. H., & Choi, C. H. (2014). The comparison of aerodynamic measures in Korean stop consonants based on phonation types. *Phonetics and Speech Sciences*, 6, 195-203.
- Dargin, T. C., & Searl, J. (2015). Semi-occluded vocal tract exercises: aerodynamic and electroglottographic measurements in singers. *Journal of Voice*, 29, 155-164.
- Dastolfo, C., Gartner-Schmidt, J., Yu, L., Carnes, O., & Gillespie, A. I. (2016). Aerodynamic outcomes of four common voice disorders: moving toward disorder-specific assessment. *Journal of Voice*, 30, 301-307.
- Delebecque, L., Pelorson, X., & Beautemps, D. (2016). Modeling of aerodynamic interaction between vocal folds and vocal tract during production of a vowel-voiceless plosive-vowel sequence. *The Journal of the Acoustical Society of America*, 139, 350-360.
- Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G\*Power 3.1: tests for correlation and regression analyses. *Behavior Research Methods*, 41, 1149-1160.
- Hirano, M. (1981). *Clinical examination of voice*. New York, NY: Springer.
- Joshi, A., & Watts, C. R. (2016). Measurement reliability of phonation quotient derived from three aerodynamic instruments. *Journal of Voice*, 30, 773.e13-773.e19.
- Kang, Y., Jo, C. W., Kim, J., & Chang, J. W. (2017). Aerodynamic parameters and the airflow regression slope (ARS) in patients with vocal polyps before and after laryngomicrosurgery according to perceptual judgment. *Clinical Archives of Communication Disorders*, 2, 15-22.
- KayPENTAX. (2010). *Phonatory Aerodynamic System (PAS) Model 6600. Software Instruction Manual*. Montvale, NJ: Author.
- Kim, H. (2012). *Neurogenic speech-language disorders*. Seoul: Sigmappress.
- Kim, J. (2014). Korean adult normative data for the KayPENTAX Phonatory Aerodynamic System Model 6600. *Phonetics and Speech Sciences*, 6, 105-117.
- Ko, H., Choi, H. S., Lim, S. E., & Choi, Y. (2015). Comparison of aerodynamic variables according to the execution methods of KayPENTAX Phonatory Aerodynamic System Model 6600. *Phonetics and Speech Sciences*, 7, 93-99.
- Korean Society of Laryngology, Phoniatrics and Logopedics. (2012). *Understanding of laryngology*. Seoul: Ilchokak.
- Lee, S. J., Choi, H. S., & Kim, H. (2018). A comparison of voice activity and participation profiles among etiological groups. *Journal of Voice*. Advanced online publication. DOI: 10.1016/j.jvoice.2018.04.016.
- Lee, S. J., Choi, H. S., Kim, H., Byeon, H. K., Lim, S. E., & Yang, M. K. (2016). Korean version of the Voice Activity and Participation Profile (K-VAPP): a validation study. *Communication Sciences & Disorders*, 21, 695-708.
- Lee, S. J., Lim, S. E., & Choi, H. S. (2017). A comparison of cepstral and spectral measures according to measurement position in a reading passage. *Communication Sciences & Disorders*, 22, 818-826.
- Lee, S. J., Pyo, H. Y., & Choi, H. S. (2018). Normative data of cepstral and spectral measures in Korean adults using vowel phonation and passage reading tasks. *Communication Sciences & Disorders*, 23, 208-217.
- Liang, F. Y., Huang, X. M., Chen, L., Huang, Y. Z., Zhang, X. Y., Su, J. H., ... & Guan, Z. (2017). Voice therapy effect on mutational falsetto patients: a vocal aerodynamic study. *Journal of Voice*, 31, 114.e1-114.e5.
- Liu, X., Zheng, Y., Tian, P., Yang, J., & Zou, H. (2015). The impact of tonsillectomy with or without adenoidectomy on voice: acoustic and aerodynamic assessments. *Journal of Voice*, 29, 346-348.
- Ma, E. P., & Yiu, E. M. (2011). *Handbook of voice assessments*. San Diego, CA: Plural Pub.
- Schaeffer, N. (2017). Pre-and poststimulation study on the phonatory aerodynamic system on participants with dysphonia. *Journal of Voice*, 31, 254.e1-254.e9.
- Zhang, Z. (2017). Compensation strategies in voice production with glottal insufficiency. *Journal of Voice*. Advanced online publication. DOI: 10.1016/j.jvoice.2017.10.002.
- Zraick, R. I., Smith-Olinde, L., & Shotts, L. L. (2012). Adult normative data for the KayPENTAX phonatory aerodynamic system model 6600. *Journal of Voice*, 26, 164-176.



## 국문초록

### 공기역학적 평가에서 음성 효율성 프로토콜의 반복 측정치 비교

이승진<sup>1</sup> · 임성은<sup>2</sup> · 임재열<sup>1</sup> · 최홍식<sup>1</sup>

<sup>1</sup>연세대학교 의과대학 이비인후과학교실 및 후두음성언어의학연구소, <sup>2</sup>강남세브란스병원 이비인후과 음성언어치료실

**배경 및 목적:** 음성 효율성 프로토콜(VOEF)의 반복 측정치 간의 차이나 반복 측정의 임상적 의의에 대해서는 충분히 규명되지 않았다. 본 연구에서는 중증도를 통제하였을 때 집단과 회차에 따라 공기역학적 측정치의 차이가 있는지 알아보고자 하였다. 또한, 첫 번째 측정치와 세 측정치의 평균치가 갖는 진단적 유용성에 차이가 있는지 살펴보고자 하였다. **방법:** 음성장애 환자군 80명과 정상 대조군 80명으로 하여금 Phonatory Aerodynamic System Model 6600 (PAS)의 VOEF를 3회씩 실시하도록 하였다. GRBAS 척도를 통해 측정 한 중증도를 통제된 상태에서 집단과 회차에 따라 평균음압(PHODB), 평균음도(MEAP), 평균성문하압(MPAP), 평균호기류율(MFPHO)의 차이가 있는지 비교하였으며, 첫 번째 측정치(MPAP<sub>1</sub>, MFPHO<sub>1</sub>)와 세 회차의 평균 측정치(MPAP<sub>Avg</sub>, MFPHO<sub>Avg</sub>)의 곡선 아래 면적(AUC)을 비교하였다. **결과:** MPAP (Wilks's  $\lambda = .991$ )와 MADV (Wilks's  $\lambda = .989$ )의 경우 회차와 집단에 따른 차이가 없었던 반면, PHODB<sub>1</sub>보다 PHODB<sub>2</sub>와 PHODB<sub>3</sub>이 더 높았으며, 환자군의 MFPHO가 대조군보다 낮았다. MPAP<sub>1</sub>와 MPAP<sub>Avg</sub>의 AUC는 .683과 .687로 차이가 없었고, MFPHO<sub>1</sub>과 MFPHO<sub>Avg</sub>의 AUC 또한 각각 .644와 .625로 차이가 없었다. **논의 및 결론:** MPAP와 MFPHO는 반복 측정에 따른 측정치와 진단적 유용성의 차이가 없어 VOEF의 1회 시행으로 충분할 것으로 보이며, 공기역학적 평가의 시간 대비 효율성을 제고할 수 있을 것으로 보인다.

**핵심어:** 공기역학적 평가, 음성 효율성, 반복 측정, 성문하압, 평균 호기류율

## 참고문헌

- 고혜주, 최홍식, 임성은, 최예린(2015). KayPENTAX Phonatory Aerodynamic System Model 6600의 수행방법에 따른 공기역학 변수 비교. *말소리와 음성과학*, 7, 93-99.
- 김재욱(2014). KayPENTAX Phonatory Aerodynamic System Model 6600을 이용한 한국 성인의 공기역학적 변수들의 정상치. *말소리와 음성과학*, 6, 105-117.
- 김향희(2012). *신경언어장애*. 서울: 시그마프레스.
- 대한후두음성언어의학회(2012). *후두음성언어의학 I: 발성의 이해와 음성치료*. 서울: 일조각.
- 이승진, 임성은, 최홍식(2017). 문단 내 위치에 따른 캡스트럼 및 스펙트럼 측정치 비교. *언어청각장애연구*, 22, 818-826.
- 이승진, 최홍식, 김향희, 변형권, 임성은, 양민교(2016). 음성 활동 및 참여 프로파일-한국판(K-VAPP): 타당성 검증 연구. *언어청각장애연구*, 21, 695-708.
- 이승진, 표화영, 최홍식(2018). 모음 발성 및 문단 읽기 과제 시 한국 성인의 캡스트럼 및 스펙트럼 측정치 정상 기준 연구. *언어청각장애연구*, 23, 208-217.
- 최성희, 최철희(2014). 한국어 파열음의 발성 유형에 따른 공기역학 측정치 비교. *말소리와 음성과학*, 6, 195-203.