

Usefulness of Glottal Inverse Filtering Analysis in Pathological Voice

Hee June Park¹, Bum Joo Shin^{2*}

¹ Dept. of Speech and Hearing Therapy, Catholic University of Pusan, Professor

² Dept. of Applied IT Engineering, Pusan National University, Professor

Purpose: In this study, we introduced Aalto Aparat, a piece of open-source software that can automatically execute glottal inverse filtering (GIF) and analyze various parameter values. We also tried to determine the usefulness of various parameters through normal and pathological voice analysis.

Methods: GIF analysis was performed on 20 normal adults and 20 with vocal nodules, and 20 with vocal cords palsy. Using Aalto Aparat, two steps required in the GIF study were inverse-filtered and converted into a glottal waveform, and the converted glottal waveform was analyzed in various ways. Analysis variables included open quotient (OQ), speed quotient (SQ), closing quotient (CLQ), amplitude quotient (AQ), and normalized amplitude quotient (NAQ), and quasi-open quotient (QQQ).

Results: By converting a general voice waveform into an LPC spectrum for converting to a glottal waveform, a glottal airflow waveform is obtained, and this waveform is integrated again to confirm that the glottal inverse filter analysis waveform is automatically analyzed without user's special manipulation. The OQ, SQ, CLQ, AQ, NAQ, and QQQ values of normal speech using GIF were analyzed to establish a reference value. In addition, it was found that the values of AQ, CLQ and OQ are useful for distinguishing between pathological voice, vocal cord paralysis, vocal cord cysts and vocal cord nodules.

Conclusions: In this study, objective values of the glottal waveforms of normal adults were obtained, and these parameter values were applied to negative diseases to determine whether there was a difference from normal voices; it was found that the GIF value is useful for pathological voice diagnosis.

Keywords: Glottal inverse filtering, open and closing quotient, speed quotient, voice analysis

Correspondence: Bum Joo Shin, PhD

E-mail: bjshin@pusan.ac.kr

Received: December 16, 2020

Revision revised: January 11, 2021

Accepted: January 29, 2021

This work was supported by research grant of Pusan National University.

ORCID

Hee June Park

<https://orcid.org/0000-0002-0522-5220>

Bum Joo Shin

<https://orcid.org/0000-0001-8638-4524>

1. 서 론

사람은 음성을 통해 자연스러운 의사소통을 하며 말소리를 위한 음성 산출은 복잡한 절차로 이루어진다. 음성은 성별, 나이, 감정, 신체적 상태 및 문화적 환경에 대한 중요한 정보를 제공한다. 이러한 음성은 신체의 기능적, 기질적, 신경학적 원인에 따라 음성장애가 올 수 있으며 잘못된 음성 사용의 장기화는 의사소통장애로 이어질 수 있다. 대표적인 음성장애로는 성대낭종, 성대용종, 성대결절, 성대마비 등이 있으며 성대의 불완전한 접촉으로 인해 쉼 목소리가 발생한다(Stemple & Hapner, 2019).

음성장애의 심한 정도를 평가할 수 있는 방법으로는 청지각적 검사, 음향학적 검사, 생리학적 검사 등이 대표적이며 이 중 청지각적 검사와 음향학적 검사의 경우에는 언어병리학에서 많은 연구

가 이루어지고 있다(Roy et al., 2013).

성대 진동에 대한 연구 방법으로는 후두스트로보스코피, 초고속 성대 촬영기, 비디오카이모그래피 등을 이용한 직접적인 방법이 있으며, 전기성문파형검사, 초음파성문파형검사, 광성문파형검사, 성문 역 여과 분석 등의 간접적인 검사 방법이 있다. 간접적인 검사 방법은 직접적인 검사방법에 비해 비싸지 않으면서 비교적 간단한 조작만으로 생리적인 상황하의 성대 진동 파형을 그래프로 할 수 있는 유용한 검사이다(Choi et al., 1994). 간접적인 검사 방법 중 성문 역 여과 분석(glottal inverse filtering: GIF)은 입술 밖으로 발산되는 소리 파형(acoustic waveform)이나 공기 흐름 파형(airflow waveform)으로부터 성도에서 변형되는 일련의 과정, 즉 성도 효과(vocal tract effect)를 역 여과(inverse filtering)시킴으로써, 성대 수준에서의 공기의 흐름을 그래프로 나타내는 방법이다(Alku, 2011).

GIF는 마이크를 통해 입력된 음성파형을 역여과하여 LPC(linear prediction coefficient) 스펙트럼으로 변환한 뒤 성문 기류 파형을 얻을 수 있다. GIF는 녹음된 음성 신호에서 성도 여과 모델을 추정한 다음 이 파형을 다시 적분(integrate)하여 성문 역 여과 분

적 파형을 구할 수 있다(Alku, 2011). GIF 분석을 통해 측정할 수 있는 다양한 파라미터들은 시간-영역(time-domain)과 주파수 영역(frequency-domain) 기반에 따라 달라진다. 시간-영역 기반 파라미터로는 시간의 흐름에 따른 성문의 개방 및 폐쇄되는 순간을 추출하여 수치로 표현하는데 대표적으로 성문 개방 지수, 속도 지수, 폐쇄 지수가 있다. 주파수-영역 기반 파라미터로는 배음 대 배음(H1-H2) 비율, 포물선 스펙트럼 기울기 등이 있다(Sundberg et al., 1999).

GIF을 분석할 수 있는 소프트웨어는 대부분 MATLAB 기반의 스크립트를 활용하는 방법이며 이는 프로그래밍 능력이 없는 일반 연구자들이 접근하기가 어려우며 이러한 점을 개선하기 위해서는 Praat와 같은 접근성과 사용방법이 간단한 프로그램이 필요하다. 현재까지 GIF 분석을 위한 단일 소프트웨어로는 DeCap, VoiceSauce, TTK Aparat 등이 대표적으로 사용되어 지고 있다. DeCap은 성문 역 여과를 통해 H1-H2, 정규화 진폭 지수 등을 분석할 수 있다(Granqvist et al., 2003). VoiceSauce는 음질 분석(jitter, shimmer, NHR 등)과 함께 성문 역 여과 분석 파라미터(H1-H2)를 분석할 수 있다(Shue et al., 2011). 마지막으로 TTK Aparat로 DeCap의 분석 파라미터에서 사용자가 세부 설정을 할 수 있도록 하여 미세 분석이 가능하도록 하였다. 세 프로그램 모두 음성분석 결과에서 강한 상관관계가 나타나 어떤 소프트웨어를 사용하든지 비슷한 결과를 얻을 수 있으며 사용 목적에 따라 적절한 소프트웨어를 사용할 수 있다. TTK Aparat의 업그레이드 버전이 Aalto Aparat로 기존 소프트웨어에서 그래픽 요소를 강화하여 사용자 친화적 인터페이스를 적용한 것이 특징이다.

GIF에 대한 선행연구를 살펴보면 파킨슨 음성과 정상 음성을 비교하여 성문 역 여과 분석이 파킨슨 음성을 구별하는데 유용하였음을 보고하였고(Corcoran et al., 2019), 후두암 진단을 위한 인공 신경망(artificial neural network) 알고리즘으로도 사용되었다(Aicha & Ezzine, 2016). 또한 인공지능 기반의 병리적 음성을 탐지하여 조기 진단을 위한 연구에도 활용되고 있다(Narendra & Alku, 2020; Zhang et al., 2018). GIF의 분석의 다양한 파라미터들이 있음에도 불구하고 임상 및 연구 목적으로 이용되기에는 많은 제한점이 있다. 입력된 음성 신호를 사용자가 LPC 스펙트럼 변환과 역 여과 필터를 위해 적분하는 과정 등 복잡한 절차를 따라야 한다는 점과 역 여과 된 성문 필터를 통해 측정할 수 있는 파라미터를 연구자나 임상가가 설정하여 파형을 수동적으로 분석해야 하는 단점이 있다. 또한 선행 연구에서 병리적 음성에 대한 파라미터들을 분석하여 인공지능 기반 조기 탐지를 위한 연구가 최근 수행되고 있으나 병리적 음성에 대한 각각의 값을 비교하여 질환별 특징을 알아보는 것이 아닌 이상 값을 확인하여 병리적 여부만 확인하고 있는 단점이 있다.

이에 본 연구에서는 GIF를 자동으로 시행하고 다양한 파라미터 값을 분석할 수 있는 오픈소스 소프트웨어인 Aalto Aparat를 소개함과 동시에 정상 음성 분석과 병리적 음성 분석을 통해 다양한 GIF 파라미터의 유용성을 알아보고자 하였다.

II. 연구 방법

1. 연구 대상

본 연구의 대상은 대학병원 이비인후과를 내원하여 후두스트로보스코프 검사를 통해 성대결절(vocal nodule), 성대낭종(vocal cyst), 일측성 성대마비(unilateral vocal cord palsy)로 진단받은 성인 환자 60명을 대상으로 하였다(Table 1). 대조군은 후두질환의 병력이 없고 후두내시경 상 특이 소견이 없었던 정상 성인 남자 20명으로 선정하였다. 각 질환군의 선정 기준은 기질적, 기능적, 신경학적 음성장애의 특성을 알아보기 위해 성대낭종, 성대결절, 성대마비를 대상으로 하였다.

Table 1. Participants' information

Category	Age <i>M (SD)</i>	Gender	<i>n</i>
Nodule	42.7 (12.6)	Male	20
Cyst	40.8 (10.8)	Male	20
Palsy	45.2 (9.4)	Male	20
Normal	41.4 (5.6)	Male	20
Total			80

2. 검사 도구

1) Aalto Aparat

Aalto Aparat는 GIF 연구를 위해 설계된 MATLAB 기반의 오픈소스 소프트웨어(Alku et al., 2017)로 웹사이트(<http://research.spa.aalto.fi/projects/aparat/>)로 접속하면 다운로드 받을 수 있으며, 특별히 MATLAB 소프트웨어를 설치하지 않아도 사용이 가능하다. Aalto Aparat는 GIF 연구에서 필요한 두 단계인 일반 음성 파형을 역 여과하여 성문 파형으로 변환하는 단계와 변환된 성문 파형을 다양한 파라미터로 분석하는 단계로 이루어져 있다. 사용자 친화적 인터페이스를 제공하고 있으며 특히 많은 양의 음성 신호를 처리하고 분석할 때 유용하다. 역 필터링 시 사용자가 특정 GIF 설정을 미세 조정할 수 있는 형태로 구현되어 있어 세밀한 분석에도 적합하다(Figure 1). 본 연구에서는 Aalto Aparat와 전기성문파형검사와의 상관관계 비교에서 포먼트 수(number of formant)

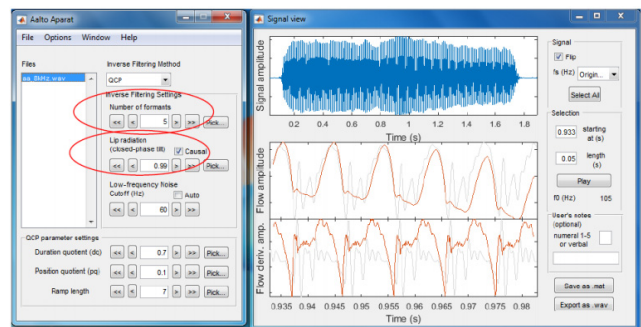


Figure 1. Two windows of Aalto Aparat: Control window left and signal view window right (Alku et al., 2017)

설정을 5개, lip radiation coefficient는 .99, low-frequency noise cutoff는 60으로 설정했을 때 가장 상관관계가 높다는 선행연구를 바탕으로 동일하게 설정하였다(Pohjalainen, 2015).

2) 분석 파라미터

성문 역 여과 후 획득한 성문 기류 파형은 다양한 매개 변수를 사용하여 자동 방식으로 분석할 수 있다(Figure 2). 분석 변수로는 개방지수(open quotient: OQ), 접촉 속도 지수(speech quotient: SQ), 폐쇄 지수(closing quotient: CLQ), 진폭 지수(amplitude quotient: AQ), 정규화 진폭 지수(normalized amplitude quotient: NAQ), 준주기 개방 지수(quasi-open quotient: QOQ)를 분석할 수 있다. 이 외에도 포물선 스펙트럼 파라미터(parabolic spectral parameter: PSP), harmonic richness factor (HRF)를 추가적으로 분석할 수 있다. 소개된 파라미터들의 자세한 내용은 Airas(2008) 연구에서 확인할 수 있다.

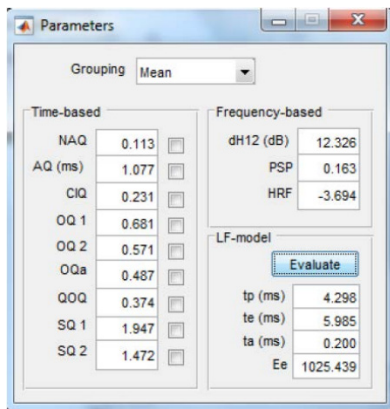


Figure 2. Results of parameterizing the glottal flow

분석 파라미터들의 구체적인 측정 방법은 Figure 3의 그림을 바탕으로 (1)~(6)과 같은 식으로 계산할 수 있다.

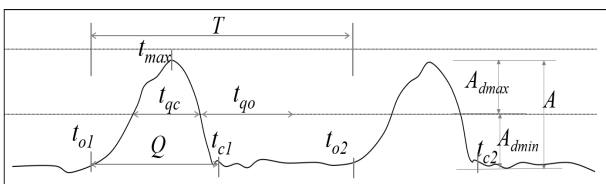


Figure 3. Time and amplitude instants used in calculating the time-domain glottal flow parameters

본 연구에서는 GIF 분석을 위해 기존 연구에서 많이 사용되고 있는 6가지 파라미터를 선정하여 사용하였으며 구체적인 파라미터들의 내용은 Table 2에 제시하였다(Price, 1989; Sapienza & Dutka, 1996; Seshadri & Yegnanarayana, 2009; Skoglund, 1998).

Table 2. GIF analysis result

Parameter	Described
OQ	The ratio between the total time of the vocal folds opening and the total time of a cycle (or period) of the glottal signal. It is inversely proportional to the intensity of the voice. The smaller it is, the higher the voice intensity. $\frac{t_{cl} - t_{o2}}{T}$
SQ	Defined as the ratio of the opening phase length to the closing phase length. $\frac{t_{max} - t_{o1}}{t_{cl} - t_{max}}$
CLQ	The ratio between the closing phase parameter (closing phase) and the total length of a glottal pulse. It is inversely proportional to the voice intensity. The smaller it is, the higher the voice intensity. $\frac{t_{o1} - t_{cl}}{T}$
QOQ	It is the relationship between the glottal signal opening at the exact instant of the oscillation and the closing time. It has been used in some works to classify emotions. $\frac{t_{qc} - t_{qo}}{T}$
AQ	The ratio between the glottal signal amplitude and the minimum value of the glottal signal derivative. It is related to the speaker phonation. $\frac{A_{dmax}}{A_{dmin}}$
NAQ	It is calculated by the ratio between the amplitude quotient (AQ) and the total time length of the glottal pulse. $\frac{AQ}{T}$

3. 실험 설계

1) 음성수집

GIF 분석을 위한 대상자의 음성 녹음은 Computerized Speech Lab(CSL, Kay Elemetrics, Montvale, NJ, USA)의 프로그램인 다중 음성 프로그램(Multi-Dimensional Voice Program: MDVP)을 사용하여 녹음하였다. 마이크를 입에서 5~10cm 간격을 두고 편안한 자세에서 대상자가 평소 사용하는 편안한 목소리로 '아' 발성을 약 5초 이상, 평균 3회 반복하여 측정하였다. 또한 Aalto Aparat 소프트웨어에서 분석이 가능하도록 WAV 파일 형태로 저장하였다.

2) 음성분석

녹음된 80명의 WAV파형은 0.5ms 단위로 각각 80개의 파일로 절삭 편집하였으며 Aalto Aparat 소프트웨어로 일괄 불러들여 매크로를 이용하여 결과 값을 획득하였다. 분석 파라미터는 NAQ, AQ, CLQ, OQ, QOQ, SQ로 하였다.

4. 결과처리

정상 음성과 병리적 음성의 GIF 분석을 통한 음성 특성을 알아보기 위하여 NAQ, AQ, CLQ, OQ, QOQ, SQ 값을 jamovi 1.6 통계 프로그램을 이용하여 독립표본 t -검정을 통해 정상음성과 병리적 음성을 각각 비교하였으며, 일원분산분석을 통해 병리적 음성의 집단 간 차이를 비교하였다(The jamovi project, 2020).

III. 연구 결과

일반적인 음성 파형을 성문 파형으로 변환하기 위해 LPC 스펙트럼으로 변환하여 성문 기류 파형을 얻고 이 파형을 다시 적분(integrate)하여 성문 역 여과 분석 파형을 사용자의 특별한 조작 없이 자동으로 분석되는 것을 Figure 4와 같이 확인할 수 있었다.

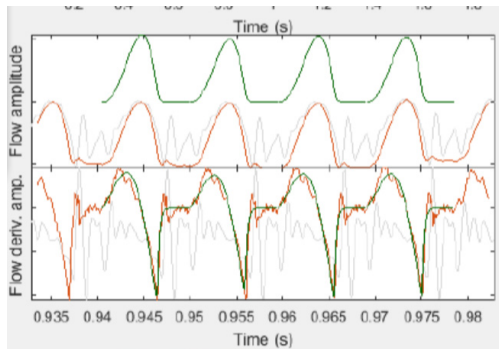


Figure 4. Signal view window after the user has selected the LPC model based parameterization

1. 정상 집단과 성대결절 집단 간의 GIF 차이

GIF를 이용한 정상 음성과 성대결절 음성의 OQ, SQ, CLQ, AQ, NAQ, QOQ 값을 분석한 결과는 Table 3과 같다. 통계분석 결과 성대결절 음성의 경우 정상 음성과 비교하여 AQ, CLQ, OQ, QOQ, SQ에서 유의하게 낮은 수치를 보였으며($p<.05$) 특히 SQ에서 정상 음성이 약 2.6에 비해 성대결절의 경우 약 1.3으로 두 배 가량 접촉 속도에 차이가 나타났다. NAQ의 경우 두 집단 간 유의한 차이는 나타나지 않았다.

Table 3. GIF difference between normal and vocal nodules

Parameter	Normal ($n=20$)	Normal ($n=20$)	t
NAQ	.103 (.018)	.121 (.052)	-1.13
AQ	.958 (.132)	.754 (.341)	2.18*
CLQ	.404 (.088)	.327 (.089)	2.30*
OQ	.900 (.081)	.791 (.085)	3.44*
QOQ	.511 (.073)	.434 (.031)	3.67*
SQ	2.606 (.517)	1.282 (.485)	6.99*

Note. Values are presented as mean (SD). NAQ=normalized amplitude quotient; AQ=amplitude quotient; CLQ=closing quotient; OQ=open quotient; QOQ=quasi-open quotient; SQ=speech quotient.

* $p<.05$

2. 정상 집단과 성대낭종 집단 간의 GIF 차이

GIF를 이용한 정상 음성과 성대낭종 음성의 OQ, SQ, CLQ, AQ, NAQ, QOQ 값을 분석한 결과는 Table 4와 같다. 통계분석 결과 정상 음성에 비해 성대낭종 음성이 AQ, CLQ, OQ, QOQ, SQ에서 유의하게 낮은 수치가 나타났으며, NAQ에서는 유의하게 높은 수치가 나타났($p<.05$).

Table 4. GIF difference between normal and vocal cyst

Parameter	Normal ($n=20$)	Cyst ($n=20$)	t
NAQ	.103 (.018)	.125 (.013)	-3.76*
AQ	.958 (.132)	1.079 (.201)	-1.88*
CLQ	.404 (.088)	.326 (.034)	3.11*
OQ	.900 (.081)	.773 (.096)	3.78*
QOQ	.511 (.073)	.430 (.061)	3.22*
SQ	2.606 (.517)	1.402 (.121)	8.45*

Note. Values are presented as mean (SD). NAQ=normalized amplitude quotient; AQ=amplitude quotient; CLQ=closing quotient; OQ=open quotient; QOQ=quasi-open quotient; SQ=speech quotient.

* $p<.05$

3. 정상 집단과 성대마비 집단 간의 GIF 차이

GIF를 이용한 정상 음성과 성대마비 음성의 OQ, SQ, CLQ, AQ, NAQ, QOQ 값을 분석한 결과는 Table 5와 같다. 통계분석 결과 CLQ, OQ, QOQ, SQ에서 성대마비 음성이 정상 음성보다 통계적으로 유의하게 낮은 수치가 나타났($p<.05$).

Table 5. GIF difference between normal and vocal cord palsy

Parameter	Normal ($n=20$)	Palsy ($n=20$)	t
NAQ	.103 (.018)	.129 (.059)	-1.60
AQ	.958 (.132)	1.038 (.500)	-0.57
CLQ	.404 (.088)	.272 (.091)	3.99*
OQ	.900 (.081)	.690 (.235)	3.16*
QOQ	.511 (.073)	.400 (.200)	1.97*
SQ	2.606 (.517)	1.537 (.715)	4.58*

Note. Values are presented as mean (SD). NAQ=normalized amplitude quotient; AQ=amplitude quotient; CLQ=closing quotient; OQ=open quotient; QOQ=quasi-open quotient; SQ=speech quotient.

* $p<.05$

4. 성대결절, 성대낭종, 성대마비 집단 간 GIF 차이

GIF를 이용한 병리적 음성의 OQ, SQ, CLQ, AQ, NAQ, QOQ 값을 분석한 결과는 Table 6과 같다. 각 파라미터 값이 집단 간 차이가 있는지 알아보기 위해 일원배치분산분석을 실시한 결과 AQ, CLQ, OQ에서 집단 간 유의한 차이가 있는 것으로 나타났다($p<.05$).

Table 6. GIF's descriptive statistics and One-way ANOVA results

Parameter	Nodule (<i>n</i> =20)	Cyst (<i>n</i> =20)	Palsy (<i>n</i> =20)	<i>F</i>
NAQ	.12 (.05)	.13 (.01)	.13 (.06)	.07
AQ	.75 (.34)	1.08 (.20)	1.04 (.50)	4.91*
CLQ	.33 (.09)	.33 (.03)	.27 (.09)	3.31*
OQ	.79 (.06)	.77 (.10)	.69 (.24)	3.21*
QOQ	.43 (.03)	.43 (.06)	.40 (.20)	.22
SQ	1.28 (.49)	1.40 (.12)	1.54 (.72)	.67

Note. Values are presented as mean (*SD*). NAQ=normalized amplitude quotient; AQ=amplitude quotient; CLQ=closing quotient; OQ=open quotient; QOQ=quasi-open quotient; SQ=speech quotient.

**p*<.05

AQ, CLQ, OQ에서 각 집단 간 세부적인 수행능력의 차이를 알아보기 위해 사후분석을 실시하였다(Table 7). AQ의 경우 성대결절이 성대낭종, 성대마비 보다 통계적으로 유의하게 높은 수치를 나타내었고(*p*<.05) 성대낭종과 성대마비에서는 유의한 차이가 나타나지 않았다. CLQ의 경우 성대마비가 성대결절, 성대낭종에 비해 유의하게 낮은 수치가 나타났으며(*p*<.05) 성대결절과 성대낭종에서는 유의한 차이가 나타나지 않았다. 마지막으로 OQ에서는 성대마비가 성대결절, 성대낭종에 비해 유의하게 낮은 수치가 나타났으며(*p*<.05) 성대결절과 성대낭종 사이에는

유의한 차이가 나타나지 않았다. 이상의 결과를 종합해 볼 때 AQ의 경우 성대결절, CLQ와 OQ는 성대마비에서 유의한 차이를 나타내는 파라미터로 판단된다.

Table 7. Comparison of Tukey Post-Hoc Test

Parameter	Group	Nodule	Cyst
AQ	Cyst	-.116*	-
	Palsy	-.086*	-.202
CLQ	Cyst	-.665	-
	Palsy	1.155*	.491*
OQ	Cyst	-.076	-
	Palsy	-.132*	-.021*

Note. AQ=amplitude quotient; CLQ=closing quotient; OQ=open quotient.

**p*<.05

집단 간 GIF 분석 결과에 대해 파라미터별로 시각화하여 비교하였으며 결과는 Figure 5와 같다. 세 집단 중 성대마비의 경우 평균값에 비해 표준편차가 높게 나타나 성대마비의 심한 정도에 따라 값의 차이가 큰 것으로 나타났다.

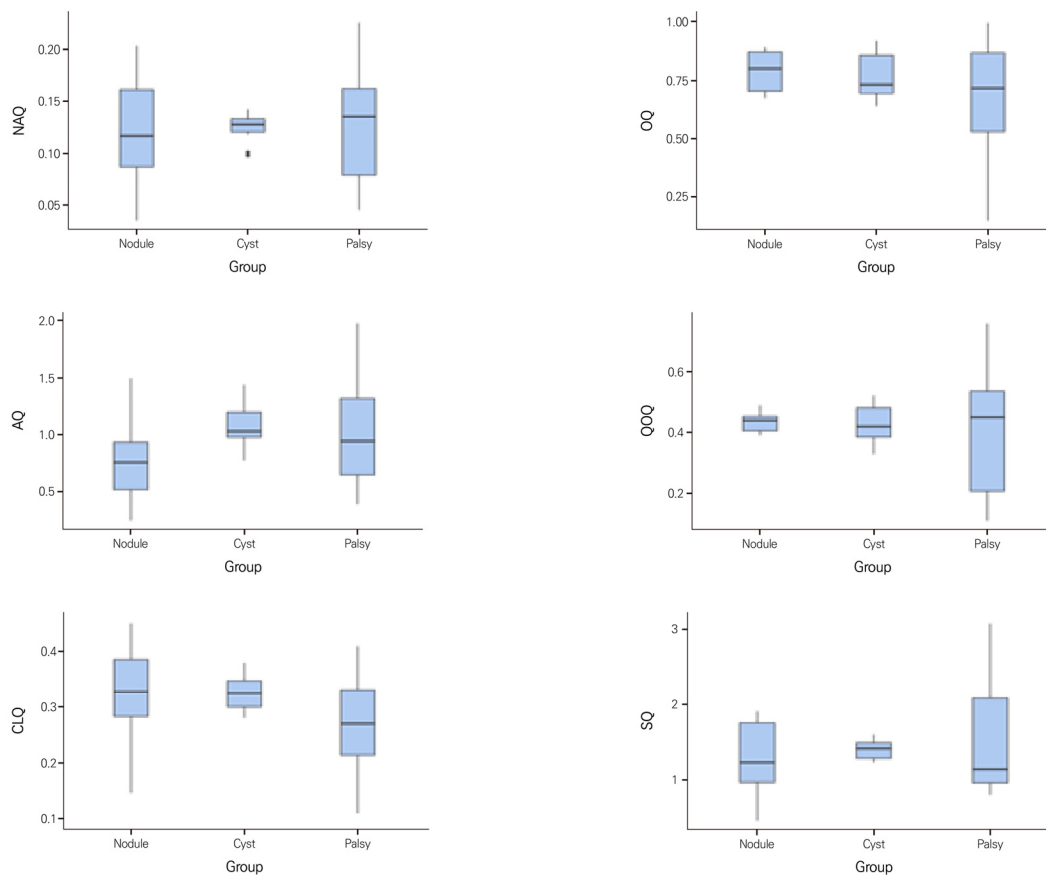


Figure 5. Comparison between groups of GIF parameters

IV. 논의 및 결론

이 연구는 새로운 GIF 변환 및 분석 도구인 Aalto Aparat를 이용하여 정상 성인의 역 성문파형의 매개변수의 값을 확인하였다. 과거 역 여과 분석을 위해서는 MATLAB이나 다양한 소프트웨어를 사용자가 프로그래밍하여 조작해야 하는 번거로움이 있었지만 본 소프트웨어를 이용하면 GIF를 자동으로 변환은 물론 OQ, SQ, CLQ, AQ, NAQ, QOQ 값을 분석할 수 있는 장점이 있다. 또한 소개된 소프트웨어는 두 가지 운영체제(Windows, Mac)에서 모두 사용할 수 있으며 성문 역 필터링에 기본 지식이 있는 사용자라면 누구나 활용할 수 있다.

본 연구에서는 정상 성인의 OQ, SQ, CLQ, AQ, NAQ, QOQ 값을 기술통계를 통해 확인 하였으며 이러한 파라미터들의 값이 성문폐쇄 부전이 있는 음성질환(성대마비, 성대결절, 성대낭종)에 적용하여 정상 음성과의 변별성이 있는지 확인해 보았다. 성대결절의 경우 AQ, CLQ, OQ, QOQ, SQ, 성대낭종의 경우 AQ, CLQ, OQ, QOQ, SQ, 성대마비의 경우 AQ, CLQ, OQ에서 정상 음성에 비해 통계적으로 유의한 차이를 나타내었다. Forero 등(2016)과 Kohler 등(2011)의 연구에서 정상음성과 성대결절, 성대마비 비교 연구에서 AQ, CLQ, OQ에서 유의한 차이를 나타낸 결과와 일치하였다. Mittal과 Sharma(2019)의 연구에서 성대결절의 경우 AQ, CLQ, OQ, SQ의 값이 정상 음성과 변별하는데 유의한 지표라고 하였으며, 성대마비 연구에서 OQ 값이 정상과 구분할 수 있고 성대마비의 회복 여부를 판단하는데 좋은 변수임을 제시한 Zagolski와 Carlson(2002)의 결과와도 일치하였다. 성대낭종의 경우 수술 전 후 전기성문파형검사에서 OQ, CLQ 값의 변화를 통해 개선 여부를 판단할 수 있다는 선행연구(Kim et al., 2019)와 비교해 볼 때 성대낭종과 정상 음성과의 비교에서 두 변수 값이 유용하다는 연구를 뒷받침한다.

병리적 음성에 대해 각 집단별 GIF 파라미터의 차이를 살펴 본 결과 AQ, CLQ, OQ에서 집단 간 유의한 차이를 보여 성대결절과 성대낭종, 성대마비를 변별하는 데 유용한 파라미터로 사용될 수 있음을 확인하였다. Bohr 등(2013)의 초고속 성대 촬영기를 이용한 병리적 음성 연구에서 CLQ와 OQ 변수는 병리적 성대진동을 구분하는 파라미터로서 유용성이 있음을 제시하였는데 본 연구에서도 유사한 결과가 나타났다. 성대진동과 관련된 다양한 파라미터를 볼 수 있는 전기성문파형 검사에서도 병리적 음성 변별에 CQ, OQ, CLQ, AQ 값이 유용하다는 선행연구들(Behrman & Baken, 1997; Deshpande & Manikandan, 2017)을 뒷받침하는 결과라 할 수 있다. GIF 분석을 통해 병리적 음성과 구체적으로 비교한 선행연구는 없지만 병리적 음성을 탐지하는데 GIF 파라미터들이 유용하다는 선행 연구를 뒷받침 하는 결과로 판단된다(Narendra & Alku, 2020; Zhang et al., 2018).

또한 각 파라미터들의 사후분석에서 OQ의 경우 정상음성과 성대결절, 성대마비 모두와 유의한 차이를 나타내어 세 집단을 변별하는데 가장 좋은 파라미터인 것으로 나타났다. 성문 역 여과를 이용한 연구와 비교는 어렵지만 생리학적 평가인 전기성문파형에

서 세 집단 간 OQ 수치가 다르다는 선행연구와 일치하였다(Yang et al., 2004). AQ, CLQ의 경우에도 정상과 성대마비, 성대결절을 구분할 수 있는 변수로 나타났으나 일반화 하기에 데이터 수가 부족한 점이 있었다.

정상음성과 병리적 음성을 구분하기 위해 실시하는 음성검사 중 후두내시경 검사가 직접 성대를 눈으로 볼 수 있어 가장 선호되는 검사 방법이지만 매번 음성치료 때 마다 치료 결과의 피드백으로 확인하는데 사용하는 데는 한계점이 있다(Mehta & Hillman, 2012). 이를 대신해 간접적으로 성대 진동을 파악할 수 있는 전기성문파형검사를 이용할 수 있으며 유용성이 확인되어 많이 사용되어 지고 있다(Holmberg et al., 1995). 하지만 전기성문파형검사 경우에도 별도의 고가인 하드웨어를 사용해야 하는 점으로 인해 일반 임상 현장에서 쉽게 접근하는데 제한점이 있다(Colton & Conture, 1990). 이에 GIF 분석은 음성 파형만으로 간접적으로 성대 진동을 파악해 볼 수 있어 임상적 유용성이 클 것으로 판단 된다.

이상의 결과를 종합해 볼 때 GIF 분석은 청지각적, 음향학적 분석과 함께 생리학적 분석의 파라미터들을 쉽게 측정할 수 있는 장점이 있었다. 기존의 생리학적 검사인 전기성문파형검사, 후두내시경, 근전도 검사와 같이 특별한 장비를 사용하지 않고 음향학적 검사 시 사용하는 음성 자료를 소프트웨어로 쉽고 빠르게 분석할 수 있다. 즉, 하나의 음성 자료만으로 청지각적 검사, 음향학적 검사, 생리학적 검사에 대한 다양한 파라미터들을 모두 측정할 수 있고 이런 데이터 종류의 증가는 향후 머신러닝과 같은 인공지능 기술에 초석이 될 것으로 판단된다.

본 연구에서는 대상자의 수가 적었던 점, 다양한 병리적 음성을 모두 분석하지 못했던 점과 GIF 세부 파라미터들의 조절을 통해 집단 간 차이를 알아보지 못한 점이 연구의 제한점으로 볼 수 있다. 향후 이러한 제한점을 개선하고 지속적인 연구를 통해 음향학적, 청지각적, 생리학적 검사의 상관관계 및 머신러닝(machine learning) 적용을 통해 병리적 음성 특징 분석을 실시하고자 한다.

Reference

- Aicha, A. B., & Ezzine, K. (2016). *Cancer larynx detection using glottal flow parameters and statistical tools*. Paper presented at 2016 International Symposium on Signal, Image, Video and Communications (ISIVC), IEEE. (pp. 65-70). doi:10.1109/ISIVC.2016.7893963
- Airas, M. (2008). TTK Aparat: An environment for voice inverse filtering and parameterization. *Logopedics Phoniatrics Vocology*, 33(1), 49-64. doi:10.1080/14015430701855333.
- Alku, P. (2011). Glottal inverse filtering analysis of human voice production: A review of estimation and parameterization methods of the glottal excitation and their applications. *Sadhana*, 36(5), 623-650. doi:10.1007/s12046-011-0041-5
- Alku, P., Pohjalainen, H., & Airaksinen, M. (2017). Aalto Aparat: A

- freely available tool for glottal inverse filtering and voice source parameterization. *Subsidia: Tools and Resources for Speech Sciences*, 21-23.
- Behrman, A., & Baken, R. J. (1997). Correlation dimension of electroglottographic data from healthy and pathologic subjects. *The Journal of the Acoustical Society of America*, 102(4), 2371-2379. doi:10.1121/1.419621
- Choi, H. S., Cho, J. I., Kim, K. M., & Park, S. S. (1994). Electroglottographic analysis of voice in normal adults. *Korean Journal of Otolaryngology*, 37(5), 1017-1025.
- Colton, R. H., & Conture, E. G. (1990). Problems and pitfalls of electroglottography. *Journal of Voice*, 4(1), 10-24. doi:10.1016/S0892-1997(05)80077-3
- Corcoran, P., Hensman, A., & Kirkpatrick, B. (2019). Glottal Flow Analysis in Parkinsonian Speech. In *BIOSIGNALS* (pp. 116-123). doi:10.5220/0007259701160123
- Deshpande, P. S., & Manikandan, M. S. (2017). Effective glottal instant detection and electroglottographic parameter extraction for automated voice pathology assessment. *IEEE Journal of Biomedical and Health Informatics*, 22(2), 398-408. doi:10.1109/JBHI.2017.2654683
- Forero, M. LA., Kohler, M., Vellasco, M. M., & Cataldo, E. (2016). Analysis and classification of voice pathologies using glottal signal parameters. *Journal of Voice*, 30(5), 549-556. doi:10.1016/j.jvoice.2015.06.010.
- Granqvist, S., Hertegård, S., Larsson, H., & Sundberg, J. (2003). Simultaneous analysis of vocal fold vibration and transglottal airflow: Exploring a new experimental setup. *Journal of Voice*, 17(3), 319-330. doi:10.1067/S0892-1997(03)00070-5
- Holmberg, E. B., Hillman, R. E., Perkell, J. S., Guiod, P. C., & Goldman, S. L. (1995). Comparisons among aerodynamic, electroglottographic, and acoustic spectral measures of female voice. *Journal of Speech, Language, and Hearing Research*, 38(6), 1212-1223. doi:10.1044/jshr.3806.1212
- Kim, J. H., Kim, M. H., Ahn, H. G., Choi, H. S., & Byeon, H. K. (2019). Clinical characteristics and management of saccular cysts: A single institute experience. *Clinical and Experimental Otorhinolaryngology*, 12(2), 212-216. doi:10.21053/ceo.2018.00808
- Kohler, M., Mendoza, L. F., Lazo, J., Vellasco, M., & Cataldo, E. (2011). Classification of voice pathologies using glottal signal parameters. *Proceedings of Brazilian Congress on Computational Intelligence (CBIC)*, 1-8.
- Mehta, D. D., & Hillman, R. E. (2012). Current role of stroboscopy in laryngeal imaging. *Current Opinion in Otolaryngology & Head and Neck Surgery*, 20(6), 429. doi:10.1097/MOO.0b013e3283585f04
- Mittal, V., & Sharma, R. K. (2019). Classification of pathological voices using glottal signal parameters. *Journal of Computational and Theoretical Nanoscience*, 16(9), 3999-4002. doi:10.1166/jctn.2019.8284
- Narendra, N. P., & Alku, P. (2020). Glottal source information for pathological voice detection. *IEEE Access*, 8, 67745-67755. doi:10.1109/ACCESS.2020.2986171
- Pohjalainen, H. (2015). *Tools for voice source analysis: Updated Aalto Aparat and a database of continuous speech with simultaneous electroglottography* (Master's thesis). Aalto University, Helsinki, Finland.
- Price, P. J. (1989). Male and female voice source characteristics: Inverse filtering results. *Speech Communication*, 8(3), 261-277. doi:10.1016/0167-6393(89)90005-8
- Roy, N., Barkmeier-Kraemer, J., Eadie, T., Sivasankar, M. P., Mehta, D., Paul, D., & Hillman, R. (2013). Evidence-based clinical voice assessment: A systematic review. *American Journal of Speech-Language Pathology*, 22(2), 212-226. doi:10.1044/1058-0360(2012/12-0014
- Sapienza, C. M., & Dutka, J. (1996). Glottal airflow characteristics of women's voice production along an aging continuum. *Journal of Speech Language and Hearing Research*, 39(2), 322-328. doi:10.1044/jshr.3902.322
- Seshadri, G., & Yegnanarayana, B. (2009). Perceived loudness of speech based on the characteristics of glottal excitation source. *The Journal of the Acoustical Society of America*, 126(4), 2061-2071. doi:10.1121/1.3203668
- Shue, Y. L., Keating, P., Vicens, C., & Yu, K. (2011). VoiceSauce: A program for voice analysis. *Proceedings of the 17th International Congress on Phonetic Sciences*, 1846-1849.
- Skoglund, J. (1998). Analysis and quantization of glottal pulse shapes. *Speech Communication*, 24(2), 133-152. doi:10.1016/S0167-6393(98)00008-9
- Stemple, J. C., & Hapner, E. R. (2019). *Voice therapy: Clinical case studies* (5th ed.). San Diego: Plural Publishing.
- Sundberg, J., Cleveland, T. F., Stone Jr, R. E., & Iwarsson, J. (1999). Voice source characteristics in six premier country singers. *Journal of Voice*, 13(2), 168-183. doi:10.1016/S0892-1997(99)80021-6
- The jamovi project. (2020). jamovi. (Version 1.6) [Computer Software]. Retrieved from <https://www.jamovi.org>.
- Yang, X., Lei, K., Gong, J., Shen, J., & Zhu, J. (2004). The clinical characteristic of electroglottography curves of pathological voice in adult. *Journal of Clinical Otorhinolaryngology*, 18(7), 430-433.
- Zagolski, O., & Carlson, E. (2002). Electroglottographic measurements of glottal function in vocal fold paralysis in women. *Clinical Otolaryngology & Allied Sciences*, 27(4), 246-253. doi:10.1046/j.1365-2273.2002.00571.x

병리적 음성에 대한 성문 역 여과 분석의 유용성

박희준¹, 신범주^{2*}

¹ 부산가톨릭대학교 언어청각치료학과 교수

² 부산대학교 IT응용공학과 교수

목적: 본 연구에서는 성문 역 여과(GIF) 분석을 자동으로 실행하고 다양한 매개 변수 값을 분석할 수 있는 오픈 소스 소프트웨어인 Aalto Aparat를 사용하여, 정상 및 병리적 음성 분석을 통해 다양한 매개 변수의 유용성을 알아보려고 하였다.

방법: 정상 성인 20명과 성대 결절 20명, 성대낭종 20명, 성대 마비 20명을 대상으로 GIF 분석을 실시하였다. Aalto Aparat를 사용하여 GIF 연구에 필요한 두 단계를 역 여과하여 성문 파형으로 변환하고 변환된 성문 파형을 다음과 같이 분석 하였다. 다양한 방법. 분석 변수로는 개방지수(OQ), 속도지수(SQ), 폐쇄지수(CLQ), 진폭지수(AQ), 정규화 된 진폭 지수(NAQ), 준개방 지수(QOQ)를 분석하여 비교하였다.

결과: 일반 음성 파형을 GIF로 분석하기 위해 LPC 스펙트럼으로 변환하여 성문 기류 파형을 얻고 이 파형을 다시 통합하여 성문 역 여과 분석 파형으로 획득했으며, 이러한 과정은 사용자의 특별한 조작 없이 자동으로 분석되는 것을 확인하였다. GIF를 이용한 일반 음성의 OQ, SQ, CLQ, AQ, NAQ, QOQ 값을 분석하여 기준 값을 제시하였다. 또한 성대결절, 성대낭종, 성대마비를 구분하는 데 AQ, CLQ, OQ 값이 유용함을 알 수 있었다.

결론: 본 연구에서는 정상 성인의 성문 파형에 대한 객관적인 값을 획득하고, 이러한 매개 변수 값을 음성 질환에 적용하여 정상 음성과 차별이 있는지 확인하였으며, GIF 값이 병리적 음성 진단에 유용함을 알 수 있었다.

교신저자: 신범주(부산대학교)

전자메일: bjshin@pusan.ac.kr

게재신청일: 2020. 12. 16

수정제출일: 2021. 01. 11

게재확정일: 2021. 01. 29

이 과정은 부산대학교 기본연구지원사업(2년)의 지원을 받아 수행된 연구임.

ORCID

박희준

<https://orcid.org/0000-0002-0522-5220>

신범주

<https://orcid.org/0000-0001-8638-4524>

검색어: 성문 역 여과 분석, 성대개방-폐쇄지수, 속도 지수, 음성분석

참 고 문 헌

- 최홍식, 조정일, 김광문, 박성수 (1994). 정상 성인에서 전기성문파형검사를 이용한 음성검사. **한이인지**, 57(5), 1017-1025.