

음성 검사법

성균관대학교 의과대학 강북삼성병원 이비인후과학교실

진 성 민

Voice Evaluation

Sung Min Jin, MD

Department of Otolaryngology, Kangbuk Samsung Hospital, School of Medicine, Sungkyunkwan University, Seoul, Korea

서 론

음성검사법의 종류를 어떻게 분류 하느냐에 대해서는 여러 가지 의견들이 있으나, 목소리를 만드는 기관 또는 만들어진 목소리에서 어떤 부분들을 알아보고자 하느냐에 따라 분류를 하는 것이 일반적이다. 따라서 음성검사를 후두의 전반적인 상태를 보고자 하는 이학적 검사, 성대점막의 움직임을 알아보는 성대진동검사, 폐로부터 만들어진 공기가 성대를 통과하면서 얼마나 효율적으로 공기들이 이용되는지를 알아보는 공기역학적 검사, 검사자가 목소리를 듣고 주관적으로 평가하는 청 지각검사, 목소리의 음향학적 요소를 분석하는 음향학적 검사 그리고 발성과 관련된 신경과 근육의 상태를 알아보는 신경생리학적 검사로 크게 분류해서 생각하면 음성검사라는 것을 보다 쉽게 이해할 수 있다. 또한 이와 같은 다양한 검사들은 단독으로 시행되기 보다는 묶음의 형태로 시행할 때 음성기능과 관련된 미세한 변화에 대하여 보다 의미 있는 개관적 자료를 얻어낼 수 있다.^{1,2)}

여기서는 외래 음성검사실에서 쉽게 검사하고 임상적

으로 유용하게 적용할 수 있는 성대진동 검사, 공기역학적 검사, 청지각 검사 및 음향학적 검사에 대하여 알아보 고자 한다.

본 론

성대진동 검사(Vibratory function measurement of vocal fold)

성대진동검사는 성대 접촉면의 움직임에 대한 정보를 얻을 수 있는 검사이다. 좋은 목소리를 만들어 내기 위한 성대의 복잡한 운동에서 건강한 성대 진동면의 상태를 유지하는 것은 매우 중요하며, 이에 대한 검사를 실시하는 것 또한 음성검사에서는 필수 요소라 할 수 있다. 일반적인 발성에서 초차도 초당 100~200회의 빠른 진동을 하는 사람의 성대 점막진동 상태는 사람의 나안으로는 관찰이 어렵기 때문에 여러 가지 기계를 이용하여 진동상태를 관찰 할 수 있다. 그 종류로는 stroboscopy, kymography, electroglottography, photoglottography 등이 있으며, 이들 중 stroboscopy는 단독으로 후두에 대한 가장 많은 임상적 정보를 제공해 줄 수 있는 검사법이어서, 임상에서도 가장 널리 사용되고 있다.^{2,3)}

교신저자 : 진성민, 110-746 서울 중로구 새문안로 29
성균관대학교 의과대학 강북삼성병원 이비인후과학교실
전화 : (02) 2001-2264 · 전송 : (02) 2001-2273
E-mail : strobojin@hanmail.net

후두스트로보스코피 검사(Laryngo-stroboscopy)

사람의 망막에 하나의 영상이 들어오면 0.2초간 머물 게 되는데, 이를 잔상(persistence of vision)이라고 하며,

0.2초 보다 짧은 시간간격으로 독립된 영상들을 연속해서 망막에 비추어주면 잔상들이 마치 서로 합쳐져 연결되어있는 것과 같이 보이게 되는데 이러한 현상을 Talbot의 법칙이라고 한다. Stroboscopy는 이러한 Talbot의 법칙을 기반으로 제작된 검사법이며, 이와 같은 원리로 성대진동의 순간 순간들을 나누어서 촬영하여 이들을 연결된 것과 같이 우리 눈에 보여 줌으로써 빠르게 진동하는 성대점막의 진동을 서서히 진동하는 것과 같은 형태로 보여주는 것이다(Fig. 1).⁴⁾

후두스트로보스코피로 성대와 성대점막의 움직임을 검사할 때 환자마다 하나의 정해진 프로토콜로 유사한 음높이와 목소리의 크기로 검사를 시행하면 각각의 결과를 의미있게 비교해서 판단할 수 있다(Table 1).⁴⁾ 정해진 프로토콜에 따라 검사를 시행하면서 관찰해야 할 항목들로는 기본진동수(fundamental frequency), 진동의 규칙성(periodicity), 진폭(amplitude), 성문폐쇄(glottic closure), 대칭성(symmetry), 점막파동(mucosal wave), 무 진동부위(non-vibrating segment) 등이 있다. 이들 중, 진동의 규칙성을 볼 때는 시간의 측면과 소리 크기 즉, 점막 진동의 크기 양쪽 측면에서 이들이 일정하게 진동을 하고 있는 지를 살펴보는 것이 중요하다. 진폭은 성대

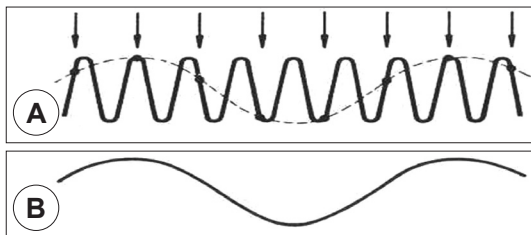


Fig. 1. Principle of stroboscopy. A : When flash intervals occur at a slightly faster rate. B : Illumination progress along each pulse and results in apparent slow motion.

Table 1. The protocol for stroboscopic examination

1. phonate /i/ at normal pitch & loudness
: sustained for at least 2 sec.
2. normal pitch and loudness and gradually getting louder.
3. normal pitch and loudness and gradually elevates pitch and then lowers the pitch.
4. to check for glottal attack
: produce a syllable chain of /i/ repetitions produced at a rapid rate.

점막의 수평움직임(extent of latero-medial)의 정도를 살펴보는 것으로 일반적으로 진성대의 보이는 부분의 1/3정도의 폭으로 점막이 진동하는 것을 정상으로 놓고, 과하거나 부족함을 판단하면 된다. 성문폐쇄는 다양한 양상으로 나타날 수 있으므로 그림을 그려서 따로 기록을 한다면 도움이 될 수 있다. 대칭성은 양측성대 점막의 움직임을 비교해볼 때 진폭은 서로 비슷한지, 점막의 수평움직임이 동일한 시간에 닫히고 열리는지를 살펴보는 항목이다. 점막파동은 진폭의 경우와 달리 점막의 상하 움직임을 살펴보는 항목이다. 마지막으로 무 진동 부위는 진폭과 점막파동이 현격하게 감소되어 있는 부위로 후두 유두종, 성대구증을 포함해서 어떠한 경우에서도 관찰할 수는 있지만, 초기 후두암의 진단과 같이 임상적 측면에서 매우 중요한 점을 시사하기도 한다. 이들 각각의 항목들은 기능성음성장애나 성대 폴립, 후두암과 같은 기질적 질환에서 각각 특징적인 변화를 보일 수 있기 때문에 이들을 잘 분석하면 환자의 상태를 보다 빠르고 정확하게 판단할 수 있다.

전기성문파 검사(Electroglottography)

성대를 직접 보면서 성대 점막 움직임을 관찰하는 방법으로 후두스트로보스코피가 가장 많이 이용되고 있는 검사라면, 간접적인 방법을 통해서 성대 점막의 움직임을 살펴보는 검사 중 가장 널리 이용되는 검사는 전기 성문파 측정법(Electroglottography, EGG)이다. 이는 후두를 통과하는 전기저항(electrical impedance)을 측정함으로써 성대가 접촉되어있는 시기(closed phase)의 상태를 알 수 있는 검사법이며, 성대 점막의 상태에 대한 정량적 해석이 가능한 검사법이다. 두 개의 전극(metal palte)을 갑상연골(thyroid alar)부위의 피부에 대고 전류를 흘려 보내서 저항(impedance)을 측정한다(Fig. 2).

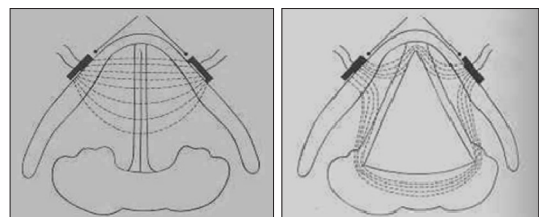


Fig. 2. High frequency current is pass through the larynx from one electrode to the other.

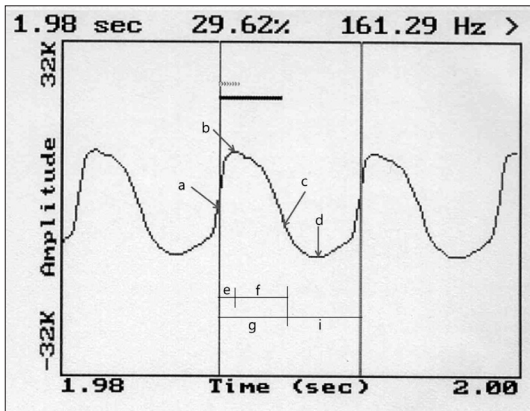


Fig. 3. Electroglottogram. a : closure of lower lip of true vocal fold, b : full adduction of true vocal fold, c : opening of lower lip of true vocal fold, d : full abduction of true vocal fold, e : closing phase, f : opening phase, g : closed phase, h : open phase.

이때 성대의 접촉면이 넓을수록 저항이 낮고 멀리 떨어질수록 높으며, 이 저항값에 따라 그래프를 얻을 수 있다. 이때 세로축은 일반적인 그래프와 달리 위로 올라갈수록 저항이 줄어드는 것으로 표시된다(Fig. 3). 이 그래프를 통해서 성대 진동 1주기(closed phase+open phase) 중에 성대가 서로 접촉되어 있는 시간(closed phase)의 비율을 측정하는 성문폐쇄율(closed quotient, CQ)를 얻을 수 있고, 성대가 접촉되어 있는 시간(closed phase) 중 성대가 닫히는 과정의 시간(closing phase)과 열리는 과정의 시간(opening phase)의 비율을 비교하는 성문접속 속도율(speed quotient, SQ)를 측정할 수 있다. 이 지표들을 통해서 성대 접착이 얼마나 효율적으로 열리고 닫히는가를 알아볼 수 있다.

공기역학적 검사(Aerodynamic study)

발성을 하는데 있어서 가장 중요한 구성요소 중의 하나가 폐로부터 나오는 적절한 공기의 흐름이다. 폐로부터 얼마나 적절한 공기가 올라오고, 이 공기를 성대에서 얼마나 효율적으로 사용하고 있는 지를 검사하는 것이 공기역학적 검사이다.

폐기능 검사 중 호흡과 관련된 측정요소는 매우 다양하지만 이들 모두가 음성분석과 연관지을 필요는 없으며, 발성 시 사용할 수 있는 공기의 양이 충분한가를 위해 측정하는 폐활량(vital capacity)을 비롯하여, 발성 동

안 이용한 공기의 양이 얼마나 되는지를 알아보기 위해서 측정하는 발성기류량(phonation volume) 등의 측정이 도움이 될 수 있다.

실제 임상에서 주로 많이 이용되는 공기역학적 검사의 측정요소 들은 최장발성지속시간(maximum phonation time, MPT), 평균호기류율(mean airflow rate), 성문하압(subglottic pressure), 발성지수(phonation quotient, PQ) 등이며, 이들은 phonatory function analyser, Aerophone II, PAS 등을 이용하여 쉽게 측정할 수 있다.⁵⁾

최장발성지속시간(Maximum phonation time)

최장발성지속시간 측정은 음성분석장비로 측정할 수도 있으나 초시계 만으로도 측정이 가능한 매우 간단하면서도 유용한 검사법이다. 최장발성지속시간만으로도 호흡 및 후두를 조절할 수 있는 능력을 복합적으로 판단할 수 있고 발성 장애의 치료 전후에 측정하여 치료 효과를 객관적으로 비교할 수 있다. 초시계를 이용한 측정법은 피험자로 하여금 호흡을 길게 들이마신 후 “아” 발성을 숨이 찰 때까지 끊이지 않고 하도록 하면서 시간을 측정하고 3회 측정을 실시하여 이 중 최대치를 사용한다. 최장발성지속시간은 짧아지는 경우 문제가 되며, 성대마비, 성대구증, 성대 폴립과 같은 성대폐쇄부전의 경우나 폐활량감소, 뇌성마비 등의 운동장애에서 값이 감소되는 것을 볼 수 있다.

평균호기율(Mean airflow rate)

발성 시 일정 시간 내에 성문을 통해 밖으로 나오는 공기의 양을 발성시 평균호기류율이라 한다. 발성기류량을 최장발성지속시간으로 나누어 그 값을 수할 수 있고 단위는 mL/sec로 표시된다. Phonatory function analyser나 aerophone II, PAS 등이 이용하여 측정하며, 지속적인 모음 발성 시 발성기능을 평가하는데 실용적인 검사법으로 공기역학검사 중 가장 의의가 큰 검사법이다. 주로 값이 커지는 경우가 문제가 되는 경우가 많은데, 성대마비가 측방으로 위치할수록 값이 커지고, 성대구증 이나 급성 후두염 등에서도 값이 커진다. 반면 연속성 발성장애에서는 다소 감소하는 경향을 보이며, 따라서 보톡스 주입술 후 효과 판정에도 이용할 수 있다.

성문하압(Subglottic pressure)

폐로부터 올라오는 호기류가 일정 압력 이상을 유지해야 효율적인 발성이 이루어진다. 따라서 성문하압(subglottic pressure)은 폐로부터 나온 공기를 발성을 위해 사용함에 있어서 성대가 얼마나 효과적으로 닫히고 있는지를 보기 위해 측정한다. 성문하압 측정은 다양한 방법으로 할 수 있지만 직접 측정 방법은 침습적이어서, 실제 임상에서는 간접 측정 방법인 기류 저지법을 많이 활용한다. 이 방법은 기기와 연결된 mouth piece를 통하여 구강내압을 측정하여 이를 성문하압을 반영하는 수치로 사용한다. 성문하압의 병적인 증가는 진행된 성문암이나 일부 기능성 음성장애에서 관찰된다.

발성지수(Phonation quotient)

발성지수는 최장발성지속시간을 측정 후 별도로 측정한 폐활량 값을 이용하여 계산하는데, 폐활량을 최장발성지속시간으로 나눈 값으로 후두염, 성대 결절, 후두암등에서 증가 하지만 평균호기율을 측정할 수 있다면 이 값은 평균호기율로 대체해도 무방하다.

청지각 검사(Perceptual analysis)

청지각적 음성평가는 객관적인 검사에 의해 평가되지 못하는 복잡한 음질의 특성을 총체적으로 평가하여 장비를 사용하지 않고도 음성장애의 정도 및 의사소통에 미치는 영향까지도 평가할 수 있다는 장점이 있어 임상에서 가장 흔하게 사용되는 평가방법이다.⁶⁾

GRBAS 척도

음성을 평가하기 위해 여러 청지각적 분석 방법들이 개발되어 왔는데, 전통적인 청지각적 음성평가 방법 중 이비인후과 임상분야에서는 GRBAS 척도가 가장 많이 알려져 있고 임상에 적용되고 있다. GRBAS 척도는 1969년 일본의 Isshiki 등이 제시한 것을 일본음성언어의학회(Japanese Society of Logopedics and Phoniatics)가 완성하고 Hirano가 1981년에 논문에 발표하면서 알려졌다. GRBAS 척도는 5개 영역으로 구성되며, G(Grade)는 음성의 전반적인 선목소리 혹은 비정상적인 음성의 정도, R(Rough)는 성대 진동의 비규칙성을 나타내는 거친 소리 정도, B(Breathy)는 성문 사이로 나오는 공기

누출에 의하여 생성되는 잡음과 관련된 바람 새는 소리 정도, A(Asthenic)는 음성의 힘이 부족하여 발생하는 약한소리 정도, 그리고 S(Strained)는 발성 시 과도한 긴장에 의해 발생하는 쥐어짜는 소리 정도를 평가하며, 각 척도는 0, 1, 2, 3의 4점 척도로 0점은 정상, 4점은 매우 나쁜 소리를 뜻한다.⁷⁾ GRBAS 척도는 검사자의 숙련성에 큰 영향을 받지 않고 검사자간의 신뢰도가 높아서 평가가 용이하며, 그 중에서도 음성의 전체적인 이상도를 평가하는 G 척도는 객관적 음성평가인 음향학적 검사와의 상관관계가 높은 것으로 보고되고 있다.^{6,7)}

CAPE-V

GRBAS 척도가 여러 가지 장점으로 널리 이용되고 있음에도 불구하고 음도, 음의 크기 등의 항목이 세분화되지 못하고 단지 음질만을 평가한다는 단점이 지적되고 있어서, 최근에는 이를 보완하여 음성언어 임상가와 음성과학자 그룹이 American Speech-Language-Hearing Association(ASHA) Special Interest Division for Voice and Voice Disorders의 후원을 받아 Consensus Auditory Perceptual Evaluation of Voice(CAPE-V ; ASHA, 2002)를 개발하였으며 최근 미국을 포함한 여러 국가에서 GRBAS 척도를 대신하여 많이 사용하고 있다.⁷⁾ CAPE-V는 검사자가 모음, 문장 및 자발화 샘플을 듣고 overall severity, roughness, breathiness, strain, pitch, loudness의 6개 주요 항목으로 나누어 평가하고, 추가로 consistent or intermittent, resonance difference, vocal fry, tremor 등의 여부 등을 평가한 후 100 mm의 선에 사선으로 표시하는 척도법이다. CAPE-V는 음질의 작은 차이 뿐 아니라 음도, 강도, 음질을 구체적으로 파악할 수 있다는 장점이 있어 임상에서 점차 사용빈도가 증가하고 있다.

음성장애지수(VHI)와 음성과 관련된 삶의 질(V-RQOL)

청지각적 평가 방법이라 할 수는 없으나, 임상에서 기계나 기구를 이용한 객관적 평가와 검사자의 청지각적 평가로 알아내기 어려운, 환자 자신만이 느끼고 호소하는 음성장애를 평가하는 데 도움이 되는 설문형식의 주관적 검사 방법들이 있다. 임상에서 많이 쓰이는 도구 들로는 음성장애지수(Voice Handicap Index, VHI)와

음성과 관련된 삶의 질(voice-Related Quality of Life, V-RQOL)이 있다. 음성장애지수(VHI)는 크게 기능(functional), 신체(physical), 감정(emotional)의 세부영역으로 나뉘며 각 영역 당 10문항씩, 각 문항은 5점 척도로 구성되어 있다. 음성과 관련된 삶의 질(V-RQOL)은 10개 항목으로 이루어져 있고 사회-심리영역과 신체영역으로 구성되어 있다. 이 검사법들은 음성장애 정도가 생활에 미치는 정도를 평가하는데 매우 유용한 주관적 평가 방법으로 이용되고 있다.⁷⁾

음향학적 검사(Acoustic analysis)

사람의 목소리를 분석하는 다양한 검사 방법들 중, 음성 신호(acoustic signal)를 컴퓨터 등을 이용하여 음파(sound wave)와 스펙트럼(spectrum) 및 스펙트로그램(spectrogram) 등으로 시각화하고 이들을 분석하는 것을 음향 음성학적 검사라 한다. 이러한 음향 음성학적 검사는 사람의 목소리에서 발성(phonation)과 조음(articulation)에 대한 음향학적 특징들을 연구하고 병적 음성에 대한 진단과 치료에 이용되기도 한다. 최근 들어 컴퓨터의 발달과 함께 목소리에 대한 다양한 음향학적 검사 프로그램들이 나와 있어 임상에서도 이 검사를 어렵지 않게 시행하고 결과를 분석할 수 있게 되었다.⁸⁾

음성분석 프로그램

사람의 목소리인 아날로그 신호를 마이크를 통하여 컴퓨터로 넣으면 아날로그 신호인 음성신호가 숫자화된 디지털 신호로 바뀌면서 컴퓨터에 저장된다. 이 디지털 신호들을 분석하여 음파로 표현하고 스펙트럼이나 스펙트로그램을 만들어 신호를 분석하고, 또한 자체의 소프트웨어 분석 프로그램을 이용하여 음성신호의 음향 지표들을 계산해내는 프로그램이 음성분석 프로그램이다. 현재 국내에 소개되어 있는 프로그램들로는 Computerized Speech Lab(CSL), Dr. Speech, Praat이 대표적이라 할 수 있다. CSL은 음성신호 분석을 위해 다양한 옵션으로 입력 및 저장이 가능하며, 다양한 종류의 스펙트럼과 스펙트로그램을 통한 음향학적 분석이 가능하다. 또한 자체적으로 음향학적 분석과 관련하여, 다양한 음향지표들을 손쉽게 측정할 수 있는 Multi-dimensional voice program(MDVP), 피검자의 최저 음

높이부터 최고 음높이 까지의 영역과 그들의 에너지량을 측정해서 발성영역을 알아볼 수 있는 Voice range profile(VRP), 구강으로 나오는 소리 에너지와 비강으로 나오는 소리 에너지의 양을 비교하여 비음도의 정도를 객관적으로 측정하고 구개인두 기능(velopharyngeal function)을 간접적으로 알 수 있는 Nasometer를 비롯하여, Visi-pitch, Sona match 등의 다양한 옵션 소프트웨어 프로그램들을 포함하고 있다. 따라서 연구, 진단, 치료 분야에 폭넓게 적용할 수 있는 장점이 있고, 고가의 음성분석 시스템이기는 하지만 국내에서 가장 널리 이용되고 있다. Dr. Speech 역시 CSL 만큼 다양하지는 않으나 거의 유사한 수준의 음향분석이 가능하고 Real analysis, Speech training, Phonetogram, Pitch master, Speech therapy 등의 옵션 프로그램들을 이용하여 연구, 진단, 치료 분야에 적용시킬 수 있다. Praat은 인터넷을 통해 무료로 받을 수 있는 음성분석 프로그램으로 많은 양의 음성 데이터를 짧은 시간에 분석해 낼 수 있고, 다양한 음향지표들을 분석할 수 있다. 그러나 음성을 여과하는 하드웨어적인 장치가 없어 주변의 녹음 환경에 따라 분석 데이터의 결과에 큰 차이를 보일 수 있고, 다른 음성분석 소프트웨어의 계산 알고리즘과 달라서 결과에 대한 객관적인 비교에 아직은 한계가 있다.⁸⁾

음성신호의 저장

최근에는 주로 컴퓨터를 이용한 디지털 신호를 분석하는 방법이 널리 사용되기 때문에, 연속된 아날로그 신호인 사람의 목소리를 컴퓨터에 저장하여 디지털 신호로 변환시킨 후 이들을 분석한다. 이러한 아날로그 신호를 디지털 신호로 변환하는 과정(digital signal processing)을 거치면서 음성신호의 정보손실을 최소화하기 위해서 반드시 알아야 할 두 가지 요소가 있는데, 표본추출률(sampling rate)과 양자화(quantization)이다.¹⁰⁾

표본추출률(Sampling rate)

표본추출률은 주파수의 정보와 관련된 요소이며 단위는 헤르츠(Hz)를 사용한다. 아날로그 신호인 원래 음향신호에서 1초당 몇 군데의 정보를 선택하여 이를 디지털 신호의 형태 즉, 숫자의 형태로 컴퓨터에 정보를 저장할 것인가를 결정하는 과정이다.

연속적인 형태의 신호인 음성신호 정보를 컴퓨터에 숫자의 형태로 입력할 때, 원신호인 음성신호의 정보를 어떻게 하면 정보의 손실이 없이 숫자로 표현할 수 있는가를 정해야 한다. 결론부터 이야기 하자면, 이를 위해서는 분석하고자 하는 최대 주파수 영역의 2배에 해당하는 값을 표본 추출률로 선택하면 된다. 예를 들어 저장된 신호를 5,000 Hz까지 분석하고 싶다면, 1초당 10,000개의 주파수와 관련된 값을 취하면 된다. 즉 표본 추출률을 10,000 Hz로 정하면 되는 것이다. 이는 분석하고자 하는 최대 주파수 영역의 2배에 해당하는 값을 표본 추출률로 선택하면 컴퓨터로 입력된 신호는 원래의 아날로그 신호와 동일한 정보를 갖게 된다는 나이퀴스트 표본추출 이론(Nyquist's sampling theory)에 따른 것으로, 일반적으로 음악 CD를 제작할 때 정하는 표본 추출률은 인간의 가청영역(약 20~20,000 Hz)의 두 배에 해당하는 44,000 Hz까지의 표본 추출률을 이용하게 된다.^{8,10,11)}

양자화(Quantization)

양자화는 진폭, 즉 소리의 강도와 관련된 요소이다. 양자화는 진폭에 관련된 정보를 얼마나 자세히 담을 것인가를 결정하는 과정으로 연속된 진폭 값을 개개의 숫자로 자르는 과정이라 할 수 있으며, 단위는 비트(bit)를 사용한다. 그런데, 양자화 값은 컴퓨터로 신호를 입력할 때 그 수치를 검사자가 직접 정하는 표본 추출률의 경우와는 달리, 일반적으로 분석기계 자체에 그 값이 정해져 있어서 검사자가 따로 그 값을 정해주지는 않는다. 보통 사람의 목소리는 8-bit 이상의 conversion을 하게 되고, 음악 CD는 16-bit 이상의 conversion을 하여, 사람의 목소리는 28단계($2^8=256$ quantization level) 이상의 값을, 음악 CD는 216단계($2^{16}=65,536$ quantization level)

이상의 값을 취하게 된다.⁸⁻¹¹⁾

스펙트럼(Spectrum)과 스펙트로그램(spectrogram)

표본 추출률, 양자화 과정과 함께 digital signal processing을 통해 컴퓨터에 음향신호가 저장되면 이들을 분석하는 과정이 필요하다. 가장 대표적인 분석방법이 스펙트럼과 스펙트로그램 분석이다.

스펙트럼

일반적인 사람의 목소리는 단순파(simple wave, sine wave)들이 합쳐져서 이루어진 복합파(complex wave)인데, 이 복합파가 어떤 단순파들이 합쳐져서 이루어진 것인가를 보여주는 방법이 스펙트럼이다.^{8,9,11-14)} 햇빛을 프리즘을 통해서 보면 빨강부터 보라까지 햇빛의 스펙트럼을 볼 수 있는데, 이와 같은 원리로 생각하면 이해하기가 쉽다. 단순파나 복합파와 같은 음파를 그래프로 표시하면 x 축은 시간, y 축은 진폭인 그래프로 표시되나, 이 음파를 분석하여 스펙트럼으로 표시하면 x 축은 주파수, y축은 진폭인 그래프로 표시된다(Fig. 4). 일반적으로 음성검사실에서 시행하는 음성신호 분석은 정확성과 신뢰성이 높은 디지털 신호 분석방법을 이용하는데, 스펙트럼 분석에 있어 디지털 신호에서 시간 함수를 주파수 함수로 변환하는 방법에는 Fourier 변환과 z 변환이 있다. Fourier 변환을 이용하여 분석하는 방법 중 고속 Fourier변환을 통해 얻어지는 FFT(fast fourier transform) 스펙트럼은 각각의 주파수 별로 에너지가 어떻게 분포되어 있는지를 쉽게 알 수 있다(Fig. 4A). z 변환을 이용하여 얻어지는 스펙트럼은 LPC(linear predictive coding) 스펙트럼이고, LPC 스펙트럼에서는 포먼트(formant)에 대한 정보를 보기가 유용하다(Fig. 4B).

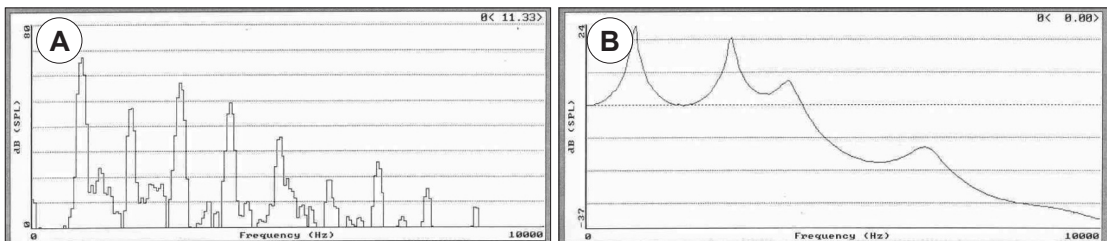


Fig. 4. Spectrum. A: Fast fourier transform (FFT). B: Linear predictive coding (LPC).

스펙트로그램

목소리는 시간에 따라 그 구성 성분의 종류와 양이 변화 한다. 따라서 스펙트럼이 시간의 개념을 생각하지 않고, 음파를 구성하고 있는 내용물의 질과 양을 보여 주는 것이라면, 스펙트럼은 시간의 개념을 추가하여 주파수와 진폭의 시간에 따른 변화를 보여주는 삼차원 적인 그림이라 할 수 있다. 따라서 스펙트럼은 x와 y의 두 축으로 구성되어 있는 반면, 스펙트로그램은 x, y, z의 세 축을 가지고 있어서 x 축을 시간, y 축을 주파수, z 축을 진폭(강도)로 표시 한다. 이때 z축은 스펙트로그램 상에서 색깔의 진하기로 표시되어 진하면 진폭(강도)이 큰 것으로, 연하면 작은 것으로 표시된다.

스펙트로그램은 음파를 분석하는 주파수 대역폭에 따라 협대역 스펙트로그램(narrow band spectrogram)과 광대역 스펙트로그램(wide band spectrogram)으로 나뉜다(Fig. 5).

협대역 스펙트로그램은 일반적으로 45 Hz 이하의 좁은 주파수 대역폭을 갖는 여과기를 이용하여 음향신호를 분석하여 얻은 그래프이며, 광대역 스펙트로그램은 300 Hz 이상의 넓은 주파수 대역폭을 갖는 여과기를 이용하여 음향신호를 분석하여 얻은 그래프이다. 각각의 스펙트로그램이 구분 되는 점은 협대역 스펙트로그램

에서는 가로줄의 가는 선들이 관찰되는 반면, 광대역 스펙트로그램에서는 세로줄이 관찰되고, 가로로 굵은 띠(포먼트)들이 관찰 된다는 점이다.^{9,12)}

협대역 스펙트로그램은 분석하고자 하는 음향신호를 구성하고 있는 각 주파수 성분을 하나하나 세밀하게 걸러내서 그 음향신호가 어떤 주파수 성분으로 구성되어 있는 지를 알아 낼 수 있다. 이와는 반대로 광대역 스펙트로그램에서는 분석하고자 하는 음향신호를 구성하고 있는 각 주파수 성분을 하나하나 세밀하게 걸러낼 수는 없지만, 특별히 어떤 주파수 대역에서 강한 음향 에너지가 몰려 있는 지를 쉽게 알아볼 수 있도록 보여 주며, 시간과 관련된 정보를 세밀하게 알 수 있다.

따라서 협대역 스펙트로그램에서는 음향신호가 가진 주파수 관련 정보를 자세하게 알 수 있지만, 시간과 관련된 정보는 잘 알 수 없는 반면, 광대역 스펙트로그램에서는 음향신호가 가진 시간 정보는 잘 알 수 있지만 주파수와 관련된 정보를 자세하게 알 수는 없다. 일반적으로 조음에 대한 연구에서는 모음의 특징과 관련된 포먼트 정보와 자음의 특징들과 관련된 시간 정보가 중요하기 때문에 협대역 스펙트로그램 보다는 광대역 스펙트로그램이 많이 사용된다.

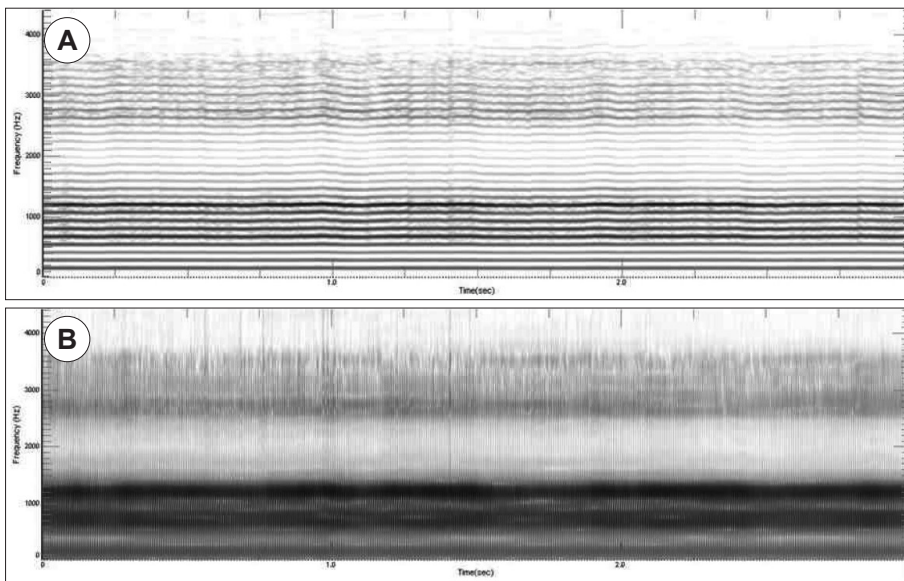


Fig. 5. Spectrogram. A : Narrow band spectrogram. B : Wide band spectrogram.

음향 지표(Acoustic parameters)

사람의 목소리는 항상 일정한 규칙성을 가지고 있는 것 만은 아니기 때문에, 소리의 파형에서 소리의 높낮이나 크기가 얼마나 불규칙하고 흔들림(perturbation)이 있는 지 그리고 잡음은 어느 정도 되는지를 객관적으로 쉽게 측정하는데 도움이 되는 음향지표들이 있다. 이러한 음향 지표는 목소리가 얼마나 거친 소리가 나는가를 객관적으로 반영 할 수 있기 때문에 임상에서는 매우 유용하게 사용되고 있다. 최근에는 앞서 언급한 음성분석프로그램을 이용함으로써 일일이 계산을 하지 않고도 음향지표의 값을 손쉽게 추출해 낼 수 있다. 일반적으로 음성의 거친 정도 즉 불규칙성을 분석하는 데는 크게 기본주파수(fundamental frequency), jitter, shimmer 그리고 잡음대 신호비(noise to harmonic ratio)의 4가지 음향지표를 측정한다.

기본주파수(Fundamental frequency)와 Jitter

앞서 언급한 바와 같이 사람의 목소리는 단순파 들이 합쳐져서 만들어진 복합파이며, 주기성을 갖는데, 기본주파수는 이러한 주기파에서 가장 낮은 주파수를 의미하며 F0의 영문 약자로 표기한다. 그런데, 일반적으로 사람의 목소리를 분석해서 추출되는 기저 주파수 값은 speaking fundamental frequency(SF0)라 하고, 이는 녹취된 목소리에서 각각의 주기 별로 측정된 기본주파수 값의 평균 값을 통해서 얻어진다.

Jitter는 주어진 기간 중 한 주기에서의 기본주파수와 그 다음 주기의 기본주파수의 가변성(variability)을 측정 한 것으로 주파수 변동률 또는 주파수 변이(frequency perturbation)라고도 한다.⁹⁾

기본주파수와 jitter는 음의 높낮이와 연관성이 있는 지표이며, 이들 값을 응용 및 변형하여 주파수의 흔들림과 관련된 부분을 relative average perturbation(RAP) 또는 pitch period perturbation(PPQ) 등 다른 종류의 다양한 음향지표들로도 표현 할 수 있다.

Shimmer

Shimmer는 주어진 기간 중 한 주기에서의 진폭과 그 다음 주기의 진폭의 가변성을 측정 한 것으로 진폭 변

동률 또는 진폭 변이(amplitude perturbation)라고도 한다. 다시 말해서 shimmer는 소리의 강도(intensity) 즉, 크기와 연관성이 있는 지표이다. Jitter의 경우와 마찬가지로 이들 값을 응용 및 변형하여 소리 강도의 흔들림과 관련된 부분을 amplitude perturbation quotient(APQ) 등 다른 종류의 다양한 음향지표들로도 표현 할 수 있다.⁹⁾

잡음대 신호비(Noise to harmonic ratio, NHR)

잡음대 신호비는 음성파형 내에서 배음(harmonics)에 대하여 잡음(noise)이 나타나는 정도를 알아보는 지표이며 불규칙하고 거친 음성일수록 잡음의 비율이 많아져서 그 값이 커지고, 이는 jitter 및 shimmer 값 등이 모두 커진 의미를 내포하고 있다. 분석 소프트웨어 또는 방법에 따라 Harmonic to noise ratio(HNR) 또는 Signal to noise ratio(SNR)로 표현할 수 있고 무엇이 분자 분모가 되느냐에 따라 적절하게 해석하면 된다.

결 론

음성에 대한 검사와 분석은 후두의 해부학적 구조와 움직임, 공기의 흐름에 대한 이해, 소리에 대한 물리학적 이해 등 다양한 부분의 지식들이 공유될 때 보다 의미 있는 결과를 얻을 수 있다. 따라서 각각의 검사와 관련된 기본적인 내용을 숙지하면서 음성검사를 시행하여야 하고, 단독으로 음성검사를 시행하기보다는 여러 가지 검사를 묶음으로 진행 함으로써 음성기능과 관련된 미세한 변화에 대하여 보다 의미 있는 개관적 자료를 얻어낼 수 있도록 해야 할 것으로 생각한다.

중심 단어 : 음성검사.

REFERENCES

- 1) Jacobson BH. *Objective voice analysis: the clinical voice laboratory*. In: Benninger MS, Jacobson BH, Johnson AF, editors. *Vocal art medicine, the care and prevention of professional voice disorders*. New York; Thieme Medical publisher;1994. p.135-52.
- 2) Sataloff RT, Heuer RJ, Emerich KA, Baroody MM, Rulnick RK, Hawkshaw M. *The clinical voice laboratory*. In: Sataloff RT, editor. *Professional voice, the science and art of clinical care*. 2nd ed. San Diego; singular publishing group;1997. p.215-43.

- 3) Sataloff RT, Spiegel JR, Hawkshaw MJ. *Stroboscovideolaryngoscopy: result and clinical value. Ann Otol Rhinol Laryngol 1991;100(9):725-7.*
- 4) Hirano M, Bless DM. *Videostroboscopic examination of the larynx. San Diego: singular publishing group;1993. p.1-21.*
- 5) 대한후두음성언어의학회. 후두음성언어의학 1. 일조각; 2012. p.369-75.
- 6) Webb AL, Carding PN, Deary IJ, MacKenzie K, Steen N, Wilson JA. *The reliability of three perceptual evaluation scales for dysphonia. Eur Arch Otorhinolaryngol 2004;261(8):429-34.*
- 7) 대한후두음성언어의학회. 후두음성언어의학 1. 일조각; 2012. p.407-16.
- 8) 대한후두음성언어의학회. 후두음성언어의학 1. 일조각; 2012. p.396-406.
- 9) Baken RJ, Orlikoff RF. *Clinical measurement of speech and voice. San Diego; Singular/Thomson Learning;2000. p.225-96.*
- 10) Kent RD, Read CR. *The acoustic analysis of speech. 2nd ed. Canada; Singular publishing group;2002. p.53-103.*
- 11) 신지영. 말소리의 이해. 제2판. 서울; 한국문화사;2001. p.125-62.
- 12) Baken RJ, Daniloff RG. *Reading in clinical spectrography and speech. 2nd ed. San Diego; Singular publishing group and Kay elemetrics Corp.;1991.*
- 13) Sataloff RT. *Voice science. San Diego; Plural Publishing; 2005. p.185-201.*
- 14) Wang SG, Roh JH, Lee BJ, Lee JC, Cheon GR, Lee SG, et al. *Development of Semiflexible Digital Laryngoscope. J Clinical Otolaryngol 2005;16(2):275-80.*