

평가자의 청지각적 음성평가와 대상자의 주관적 음성평가 비교

김재옥^{a,§} · 최홍식^b

^a강남대학교 교육대학원 언어치료교육전공, ^b연세대학교 의과대학 이비인후과학교실 음성언어의학연구소

[§] 교신저자

김재옥
강남대학교 교육대학원
언어치료교육전공 교수
경기도 용인시 기흥구 구갈동 111
e-mail: jaeock@gmail.com
tel.: 031-280-3221

배경 및 목적: 음성장애를 평가할 때 객관적 평가와 청지각적 평가가 주로 사용된다. 그러나 이들 뿐만 아니라 음성장애를 경험하는 대상자들의 주관적인 평가가 요구된다. 이에 본 연구에서는 청지각적 평가와 주관적 평가 간의 상관성을 파악하고자 한다. **방법:** 음성장애로 진단 받은 120명과 정상인 120명을 대상으로 한국어판 음성장애지수(Korean Voice Handicap Index: K-VHI, 이하 K-VHI)와 한국어판 음성과 관련된 삶의 질(Korean Voice-Related Quality of Life: K-VRQOL, 이하 K-VRQOL)을 작성하게 하였으며, 숙련된 음성언어치료사와 이비인후과 전문의가 GRBAS 척도로 평가하였다. **결과:** GRBAS 척도 중 G 척도와 K-VHI 및 K-VRQOL의 총점수와 세부영역별 점수 간에는 약한 상관관계를 보였으나, G 척도에 따른 음성장애 중증도가 높아질수록 K-VHI의 영역별 점수는 높았으며, K-VRQOL의 영역별 점수는 낮았다. 또한 진단별 분류에 따라 K-VHI의 신체 점수를 제외한 K-VHI의 모든 영역별 점수와 K-VRQOL의 모든 영역별 점수 간에 통계적으로 유의미한 차이를 보였다. **논의 및 결론:** 청지각적 평가와 주관적 평가 간에는 유의미한 상관관계가 없어 평가자가 평가한 음성장애의 정도와 대상자가 지각하고 있는 음성장애의 정도가 다를 수 있었다. 그러나 청지각적 평가 분류 및 진단별 분류에 따라 주관적 평가의 차이가 있었다. 이에, 음성장애를 평가할 때 대상자의 주관적 평가를 함께 포함하는 다양한 측면의 음성평가가 요구된다. 『언어청각장애연구』, 2009;14;223-235.

핵심어: 음성장애, 청지각적 평가, 주관적 평가, GRBAS

I. 서론

음성장애는 크게 세 영역별로 평가될 수 있다. 첫째, 객관적인 자료 수집을 위한 평가방법으로 음향학적검사(acoustic analysis), 공기역학적검사(aerodynamic analysis), 전기성문파형검사(electroglottography)나 후두스트로보스코피(laryngeal stroboscopy)와 같은 후두내시경(laryngeal endoscopy) 등이 있다. 객관적 평가방법은 음성장애의 정도를 양적으로 수치화할 수 있어 음성장애의 진단을 내리는데 구체적 자료를 제시해 줄 수 있고, 음성치료, 약물치료, 혹은 수술치료와 같은 치료 전과 후의 변화를 기록할 수 있다. 그러나 이들은 측정기구나 측정자 혹은 측정 방식에 따라 수치가 변화할 수 있고, 복잡한 음질의 특성을 총체적으로 파악하기에는 제한이 있다. 둘째, GRBAS

척도(Hirano, 1981) 또는 Consensus Auditory Perceptual Evaluation of Voice (이하 CAPE-V) (American Speech-Language Hearing Association, 2002)와 같은 평가자에 의한 청지각적 평가방법을 사용하여 대상자의 음질과 음성장애의 중증도를 평가할 수 있다. 이 방법은 객관적인 검사에 의해 평가되지 못하는 복잡한 음질의 특성을 총체적이고 융통성있게 평가할 수 있으며, 장비를 사용하지 않고도 평가할 수 있다는 장점이 있어 음성장애의 정도를 평가하는 “정석(gold standard)”이라고 여겨진다(Ma & Yiu, 2006). 그러나 사용되는 척도의 유형이나 음질 등에 영향을 받을 수 있으며, 평가자의 숙련도가 낮을 경우 평가자간의 신뢰도가 낮을 수 있고, 평가자의 주관적인 측면이 개입되는 단점이 있다(Ma & Yiu, 2006). 셋째, 음성장애를 경험하는 대상자가 자신의 음성과

* 이 논문은 2007년 정부(교육인적자원부)의 재원으로 한국학술진흥재단의 지원을 받아 수행된 연구임(KRF-2007-355-E00012).

관련된 심리적인 측면을 평가할 수 있는데, 일반적으로 많이 사용되는 평가도구로 음성장애지수(Voice Handicap Index: VHI, 이하 VHI) (Jacobson et al., 1997), 음성과 관련된 삶의 질(Voice-Related Quality of Life: V-RQOL, 이하 V-RQOL) (Hogikyan & Sethuraman, 1999), 아이오와 환자음성색인(Iowa Patient's Voice Index: IPVI, 이하 IPVI) (Karnell et al., 2007), 음성증상척도(Voice Symptom Scale: VoiSS) (Deary et al., 2003), 음성평가척도(Voice Rating Scale: VRS) (Wingate et al., 2007) 등이 있다. 대상자의 주관적인 평가방법은 객관적으로 평가된 음성장애의 정도가 비슷하더라도 각 대상자가 경험하는 음성장애의 정도를 주관적으로 평가할 수 있으므로 치료의 만족도를 평가하기에 적합하다. 특히 전문적으로 음성을 사용하는 사람들의 경우 음성에 대한 의존도가 높으며 음성 변화에 민감하기 때문에 이들의 주관적 평가는 매우 중요하다. 이러한 대상자 중심의 평가를 통해 대상자가 경험하는 주관적인 음성장애의 정도가 전반적인 음성평가에 반영이 될 때, 음성이 대상자의 일상생활에 미치는 영향이나 중재에 대한 효율성을 개별화하여 평가될 수 있다는 장점이 있다.

앞에서 언급한 바와 같이 평가자 중심의 청지각적 평가로는, 1975년 일본 음성언어의학회에서 제안한 GRBAS 척도(Hirano, 1981)와 2002년 미국언어병리학회의 Special Interest Division #3 for Voice and Voice Disorders에서 제안한 CAPE-V 척도가 많이 사용되고 있다. GRBAS 척도는 음성의 전반적인 쇠소리의 정도(Grade: G), 거친 소리의 정도(Rough: R), 바람새는 소리의 정도(Breathy: B), 약한 소리의 정도(Asthenic: A) 그리고 긴장과 쥐어짜는 소리의 정도(Strained: S)를 평가하며, 각 척도는 0, 1, 2, 3의 4점 Likert 척도로 0점은 정상, 3점은 매우 나쁜 소리를 뜻한다. GRBAS 척도 중에서도 음성의 전체적인 이상도를 평가하는 G 척도는 음향학적 검사의 일종인 Multi-Dimensional Voice Program (MDVP, Kay Elemetrics, Lincoln Park, NJ, 이하 MDVP)의 변수들과의 상관관계가 높다(표화영 외, 1999; De Bodt et al., 1997). 또한 G 척도는 평가자의 숙련성에 큰 영향을 받지 않고 평가자간의 신뢰도가 높아 평가가 용이하다(Anders et al., 1988). 그러나 GRBAS는 음도(pitch), 음의 강도(loudness)는 평가하지 못하고 단지 음질(quality)만을 평가한다는 단점이 지적되고 있

다(Bhuta, Patrick & Garnett, 2004). 이에 반해, CAPVE-V 척도는 음도, 음의 강도, 음질을 모두 평가할 수 있는 항목들이 추가되어 있으며, 시각적 아날로그 척도(visual analog scale)로 구성되어 있어 Likert 척도의 단점을 보완하였으나, 평가자간 신뢰도가 Likert 척도에 비해 다소 낮은 경향을 보인다(Wingate et al., 2007).

음성장애가 생활에 미치는 정도를 파악하는 대상자 중심의 주관적 자가보고 형식의 평가도구들 중 가장 널리 사용되고 있는 것으로, VHI와 V-RQOL이 있으며, 여러 국가에서 이들을 자국어로 번안하여 사용하고 있다. 현재 국내에서도 한국어로 번안한 후 신뢰도와 타당도를 검증한 한국어판 음성장애지수(Korean Voice Handicap Index: K-VHI, 이하 K-VHI) (김재옥 외, 2007a)와 한국어판 음성과 관련된 삶의 질(Korean Voice Related Quality of Life: K-VRQOL, 이하 K-VRQOL) (김재옥 외, 2007a)이 사용되고 있다. K-VHI는 기능(functional: F), 신체(physical: P) 그리고 감정(emotion: E)의 세 가지 세부영역으로 구분되고, 기능 영역의 문항들은 음성장애가 일상생활에 미치는 영향을 평가하며, 신체 영역의 문항들은 음성장애로 초래되는 음성 산출 기관의 불편감이나 음성 산출의 특성과 관련이 있고, 감정 영역의 문항들은 음성장애에 반응하는 대상자의 감정에 대하여 평가한다. 각 문항은 5점 척도로 0점은 “거의 그렇지 않다”이고, 4점은 “항상 그렇다”로 각 영역별로 10문항씩 총 30문항으로 구성되어 있으며, 총점수의 범위가 0점에서 120점으로 원점수가 높을수록 음성으로 초래되는 장애가 높다고 평가된다. K-VRQOL은 음성이 삶의 질에 미치는 영향을 측정하는 것으로 10항목으로 이루어져 있다. 이는 두 세부영역, 사회-심리 영역(social-emotional: SE)과 신체기능 영역(physical function: PF)으로 구성된다. K-VRQOL은 총점수 및 세부 영역의 점수에 따라 0점에서 100점으로 표준화하여 0점은 음성장애로 인한 삶의 질이 매우 나쁜 것을 뜻하고, 100점은 음성장애가 삶의 질에 영향을 미치지 않음을 뜻한다.

음성은 다차원적이기 때문에 하나의 평가방법으로는 정확한 진단과 중재를 할 수 없다. 이에 다양한 평가방법을 사용하여 종합적으로 분석할 필요가 있다(Bhuta, Patrick & Garnett, 2004; Ma & Yiu, 2006). 일반적으로 평가자에 의해 정상 음성으로 판정받아도 대상자 자신은 음성산출의 어려움이나 불편감을 호소

할 수 있으며, 대상자가 자신의 음성이 정상이라고 판단할지라도 평가자가 평가하기에는 음성장애를 지니고 있다고 할 수 있다(김재욱 외, 2007b). 선행연구에 의하면, 음향학적 분석이나 공기역학적 검사를 통한 객관적 평가와 자가보고를 통한 대상자의 주관적 평가 간의 상관관계는 거의 없어 객관적 검사결과와 대상자 자신의 음성에 대한 평가결과가 항상 일치하는 것은 아니라고 한다(Hsiung, Pai & Wang, 2002; Wheeler, Collins & Sapienza, 2005; Woisard et al., 2007). 그러나 객관적 평가와 청지각적 평가 간의 상관성을 조사한 연구(표화영 외, 1999; Bhuta, Patrick & Garnett, 2004; Hakkesteegt et al., 2008; Kreiman et al., 1993)에 의하면 객관적 평가와 평가자의 청지각적 평가는 모든 영역에서는 아니지만, 많은 부분에서 중간(moderate) 이상의 상관관계를 보였다. 그렇다면, 평가자 중심의 주관적 음성검사인 청지각적 평가와 대상자 중심의 주관적 평가 사이에는 어떠한 상관관계가 있을까? 이를 분석함으로써 객관적 평가와 청지각적 평가가 대상자가 지각하는 음성과 관련된 불편감이나 장애 정도를 얼마나 반영할 수 있는지를 파악할 수 있을 것이다. 이러한 선행연구로 유일하게 Karnell et al. (2007)의 연구가 있다. 이 연구에서는 GRBAS 척도 중 G 척도와 V-RQOL 및 IPVI와의 상관관계를 살펴보았으며, 그 결과 이들 간의 상관관계가 낮음을 밝혔다. 그러나 주관적 평가로써 가장 많이 사용되고 있는 VHI와의 상관관계는 규명하지 않았으며, 영어에 근간한 연구로 한국어로 변환된 평가 도구를 이용하여 이들의 상관관계를 조사할 필요가 있다. 이와 함께 평가자에 의해 평가된 음성장애의 중증도에 따라 주관적 평가가 어떻게 다른지를 분석함으로써 청지각적 평가와 주관적 평가 간의 관계를 좀 더 명확하게 파악할 수 있을 것이다. 또한 청지각적으로는 정상 음성을 지니고 있으나 음성 사용의 어려움을 겪는 사람들이 느끼는 불편감 정도가 정상인들과 얼마나 그리고 어떻게 차이가 있는지를 조사함으로써, 음성을 총체적으로 평가해야 하는 타당성을 살펴볼 수 있을 것이다. 더불어 음성장애의 진단별 분류에 따라 대상자들이 실제로 경험하는 음성장애의 정도의 차이가 있는지를 평가할 필요가 있다. 각기 다른 음성장애를 경험하는 대상자가 자신의 진단명을 알지 못한다 하더라도 평소 음성을 사용하는데 있어 느끼는 불편감은 음성장애의 종류에 따라 다를 것이다.

이에 본 연구는 평가자의 청지각적 평가방법의 하

나인 GRBAS 척도와 주관적 평가방법인 K-VHI와 K-VRQOL과의 상관성 정도와 그 유의미성을 분석하고, GRBAS 척도 등급에 따른 K-VHI 및 K-VRQOL의 차이를 비교하며, 음성장애군과 정상음성군의 주관적 평가의 차이를 비교하고자 한다. 또한 다양한 음성장애의 진단별 분류에 따라 주관적 평가의 차이가 있는지를 살펴보고자 한다.

II. 연구 방법

1. 연구 대상

실험군인 음성장애군과 대조군인 정상음성군은 모두 주어진 설문지를 읽고 이해하며 평가가 가능한 자였으며, 전문적 음성 사용자(가수, 성악가, 방송인, 배우, 교사, 목회자, 영업직 종사자)는 제외되었다. 일반적으로 음성을 전문적으로 사용하는 사람의 경우 자신의 음성에 대한 의존도가 높아 음성을 전문적으로 사용하지 않는 사람들보다 음성문제에 대해 민감한 반응을 보여 높은 음성장애지수를 나타내기 때문이다(Thomas et al., 2007).

가. 음성장애군

음성장애군은 2007년 9월부터 2008년 2월까지 음성문제를 주호소로 영동세브란스병원 이비인후과 음성클리닉을 방문하여 이비인후과 전문의에 의해 후두스트로보스코피를 시행하고 음성언어치료사에 의해 음향학적 검사, 공기역학적 검사 등의 음성정밀검사를 실시한 결과, 음성장애로 진단받은 120명(남성 53명, 여성 67명)을 대상으로 하였다. 이들의 평균연령은 37.98세로 연령범위는 19세에서 59세였다.

일반적으로 음성장애를 구조적, 기능적, 신경학적 장애로 구분한다. 그러나 성대결절이나 성대용종 혹은 라인케씨 부종과 같은 질환은 기능적 장애가 구조적 장애를 일으킨 것으로 구조와 기능적 장애를 구분하는 것이 어렵다(Colton & Casper, 1996). 또한 양성병변의 유무와 신경학적 장애에 따라 음성을 사용할 때 경험하는 불편감의 정도가 다를 것이라 여겨져 음성장애군의 진단명을 기준으로 크게 다음과 같이 세 집단으로 나누었다. 양성병변의 유무에 따라 성대에 양성병변이 있는 구조적-기능적 장애 집단(성대결절, 성대용종, 성대낭종, 라인케씨 부종, 유두종, 육아종,

성대백반증)과 성대에 양성병변이 없는 구조적-기능적 장애 집단(성대구, 근긴장성 발성장애, 변성 발성장애, 후두염, 남성음) 그리고 신경학적 장애 집단(성대마비, 경련성 발성장애)이었다(박영학 외, 2005; 엄재욱 외, 2007). 이들은 각각 58사례, 30사례 그리고 32사례였다.

음성장애군 대상자들은 음성치료, 약물치료, 혹은 수술치료 등의 치료를 시행하지 않은 상태에서 본 연구에 참여하였다.

나. 정상음성군

정상음성군은 대상자 스스로가 최근 3개월 이내에 음성장애를 경험하지 않았다고 보고하며, 자발화를 산출하는 동안 연구자가 청지각적으로 평가하여 GRBAS 척도 중에서 음성의 전반적인 쉼 소리 정도를 나타내는 G 척도가 0점(G0)으로 평가된 정상인 120명(남성 47명, 여성 73명)을 대상으로 하였다. 대조군의 평균연령은 32.01세로 연령범위는 20세에서 59세였다.

2. 연구 방법

가. GRBAS 평가

음성장애군이 3초간 /a/ 모음 연장발성과 '가을문단'을 읽는 동안 1명의 숙련된 음성언어치료사와 1명의 이비인후과 전문의가 개별적으로 GRBAS 척도를 평가하였다. 각 척도의 등급은 0, 1, 2, 3의 4점 척도를 기준으로 하였으며, 본 연구에서는 음성언어치료사의 척도를 사용하였으며, 자료 수집이 끝난 후 녹음된 음성장애군의 /a/ 모음 연장발성과 '가을문단' 자료 중 15% (18명)를 무작위로 추출하여 재생하면서 GRBAS 척도를 평가하였다.

나. K-VHI와 K-VQOL 작성

음성장애군과 정상음성군은 K-VHI와 K-VRQOL을 작성하였다. 대상자들이 두 설문지를 작성하기 전에 대상자 스스로가 음성으로 인해 경험하고 생각하는 것을 솔직하게 평가하도록 권유하였으며, 문항을 명확하게 이해하지 못할 경우에 음성언어치료사가 문항의 해석을 도와주었다. 각 대상자가 두 설문지를 작성하는데 소요된 시간은 평균적으로 5~10분 정도였다.

다. 통계 분석

GRBAS 척도에서 R 척도와 B 척도는 평가자의 숙련도가 매우 높아야 평가의 신뢰도가 높아지고, R 척도와 B 척도가 함께 G 척도를 구성한다(Hakkesteegt et al., 2008). 따라서 본 연구에서의 두 평가자 간의 신뢰도는 음성장애군 120명의 15%에 해당하는 18명의 G 척도를 무작위로 추출하여 이들의 상관관계를 산출하였으며, 평가자내 신뢰도는 음성언어치료사가 음성정밀평가 시 평정한 G 척도와 녹음한 음성장애군의 자료 중 무작위로 추출한 15% 자료에서 평정한 G 척도 간의 상관관계를 산출하였다. 평가자간 신뢰도와 평가자내 신뢰도 분석을 위해 Spearman r correlation을 시행하였다. G에 대한 두 평가자 간의 신뢰도는 통계적으로 유의미하게 높은 상관관계($r = .91, p < .01$)를 보였으며, 평가자내 신뢰도도 유의미하게 높은 상관관계($r = .96, p < .01$)를 보였다.

GRBAS 척도의 각 척도와 K-VHI의 총점수(Total)와 세부영역별 점수(F, P, E) 및 K-VRQOL의 표준화된 총점수(Sum)와 세부영역별 점수(SE, PF)의 상관관계는 Pearson r correlation으로 측정하였다. 또한, GRBAS 척도와 두 평가도구의 어떠한 문항들이 상관관계가 있는지를 평가하기 위하여 Pearson r correlation을 시행하였다.

음성장애군의 G 척도에 따른 K-VHI의 Total, F, P, E와 K-VRQOL의 Sum, SE, PF 점수의 차이는 G 척도별 분류에 해당하는 대상자수의 차이가 커서 비모수검정인 Kruskal-Wallis test로 비교하였다. 더불어 음성장애군과 정상음성군의 K-VHI의 Total, F, P, E와 K-VRQOL의 Sum, SE, PF 점수는 각 독립표본 t 검정(independent samples t -test)으로 비교하였으며, 음성장애군의 G 척도가 0점(G0)인 집단과 정상음성군(G0)의 차이는 두 군 간의 대상자수의 차이가 커서 비모수검정인 Mann-Whitney U test를 시행하여 비교하였다. 진단별 분류에 따른 K-VHI의 Total, F, P, E와 K-VRQOL의 Sum, SE, PF 점수의 차이는 G 척도를 공변량으로 하는 One-way ANCOVA를 사용하였으며, 진단별 분류에 따라 차이가 있을 경우에는 Bonferroni 사후검정을 실시하였다. 모든 분석결과는 SPSS 12.0™ 프로그램(SPSS Inc., Chicago, IL)을 이용하여 통계적으로 처리하였다.

Ⅲ. 연구 결과

1. 진단별 분류에 따른 G 척도별 분포

음성장애군의 진단별 분류에 따른 G 척도를 분석한 결과는 <표 - 1>과 같으며, G0인 경우는 11명, G1인 경우는 56명, G2인 경우는 43명, G3인 경우는 10명이었다.

<표 - 1> 진단별 분류에 따른 G 척도별 분포

대상자	진단별 분류	G				
		G0 (N)	G1 (N)	G2 (N)	G3 (N)	총 (N)
음성 장애군	양성병변이 있는 구조적-기능적 장애 집단	5	26	26	1	58
	양성 병변이 없는 구조적-기능적 장애 집단	4	15	6	5	30
	신경학적 장애 집단	2	15	11	4	32
정상 음성군		120	0	0	0	120
총 인원수		131	56	43	10	240

2. 음성장애군의 GRBAS 척도와 K-VHI 및 K-VRQOL 간의 상관관계

GRBAS 척도와 K-VHI의 총점수 및 세부영역별 점수 간의 상관계수의 범위는 -.034 ~ .320이었으며, GRBAS 척도와 K-VRQOL의 표준화된 총점수 및 세부영역별 점수 간의 상관계수의 범위는 -.320 ~ .308이었다(<표 - 2>). 이 중에서 G 척도와 K-VHI간의 상관계수의 범위는 .267 ~ .320이었고, G 척도와 K-VRQOL 간의 상관계수의 범위는 -.320 ~ -.244였다.

GRBAS 척도의 각 척도별 상관계수의 범위는 -.281 ~ .584였으며 S 척도를 제외하고 모두 통계적으로 유의미하였다. K-VHI 총점수와 세부영역별 점수의 상관계수의 범위는 .725 ~ .931이었고, K-VRQOL 총점수와 세부영역별 점수의 상관계수의 범위는 .855 ~ .931로, 모든 상관관계는 통계적으로 유의미하게 높았다. 또한 K-VHI와 K-VRQOL 간의 상관관계는 K-VHI는 점수가 높을수록 그리고 K-VRQOL은 점수가 낮을수록 음질이 더 나쁨을 나타내므로, 모두 부적관계를 보였으며 상관계수의 범위는 -.878 ~ -.599이었다.

K-VHI의 30문항들과 GRBAS 척도 간의 상관계수의 범위는 -.118 ~ .364였으며, K-VRQOL의 10문항과 GRBAS 척도 간의 상관계수의 범위는 -.102 ~ .369였다.

<표 - 2> 음성장애군의 GRBAS 척도와 K-VHI 및 K-VRQOL 간의 상관관계

	GRBAS				K-VHI				K-VRQOL			
	R ^{b)}	B ^{c)}	A ^{d)}	S ^{e)}	Total ^{f)}	F ^{g)}	P ^{h)}	E ⁱ⁾	Sum ^{j)}	SE ^{k)}	PF ^{l)}	
GRBAS	G ^{a)}	.453**	.584**	.402**	.212*	.320**	.320**	.305**	.267**	-.309**	-.244**	-.320**
	R		.087	-.281**	-.130	.031	-.034	.051	.054	-.029	.018	-.016
	B			.564**	-.054	.164	.219*	.165	.094	-.161	-.130	-.206*
	A				-.039	.220*	.274**	.209*	.150	-.191*	-.162	-.248**
	S					.228*	.214*	.121	.284**	-.245**	.308**	-.258**
K-VHI	Total					.931**	.903**	.927**		-.878**	-.791**	-.818**
	F						.799**	.797**		-.838**	-.732**	-.807**
	P							.725**		-.745**	-.599**	-.730**
	E									-.834**	-.838**	-.741**
K-VRQOL	Sum									.906**	.931**	
	SE											.855**

* $p < .05$, ** $p < .01$

a) G: Grade, 전반적인 신 소리의 정도; b) R: Rough, 거친 소리의 정도; c) B: Breathly, 바람 세는 소리의 정도; d) A: Asthenic, 약한 소리의 정도; e) S: Strained, 긴장과 쥐어짜는 소리의 정도; f) Total: K-VHI의 총점수; g) F: functional, 기능 점수; h) P: physical, 신체 점수; i) E: emotional, 정서 점수; j) Sum: K-VRQOL의 표준화된 총점수; k) SE: social-emotional, 사회-심리영역 점수; l) PF: physical-functional, 신체기능 영역 점수

<표 - 3> 음성장애군의 G 척도에 따른 K-VHI 및 K-VRQOL의 비교

		음성장애군 G 척도				χ^2
		0	1	2	3	
		중위수 (1사분위수, 3사분위수)	중위수 (1사분위수, 3사분위수)	중위수 (1사분위수, 3사분위수)	중위수 (1사분위수, 3사분위수)	
K-VHI	Total ^{a)}	26 (12, 79)	47 (34, 68)	51 (37, 73)	78 (65, 105)	14.26**
	F ^{b)}	12 (3, 21)	15 (10, 22)	17 (12, 23)	27 (22, 32)	15.29**
	P ^{c)}	10 (7, 26)	21 (17, 27)	22 (16, 28)	33 (27, 33)	16.66**
	E ^{d)}	13 (1, 29)	12 (7, 21)	14 (7, 23)	23 (10, 37)	9.55*
	Sum ^{e)}	80.00 (57.50, 100.00)	72.50 (50.63, 87.50)	62.50 (42.50, 82.50)	42.50 (23.75, 50.63)	15.03**
K-VRQOL	SE ^{f)}	81.25 (50.00, 100.00)	75.00 (51.56, 93.75)	75.00 (37.50, 87.50)	43.75 (20.31, 50.00)	11.18*
	PF ^{g)}	87.50 (50.00, 100.00)	66.67 (45.83, 88.33)	58.33 (41.67, 79.17)	33.33 (25.00, 51.04)	15.26**
N		11	56	43	10	

* $p < .05$, ** $p < .01$

a) Total: K-VHI의 총점수; b) F: functional, 기능 점수; c) P: physical, 신체 점수; d) E: emotional, 정서 점수; e) Sum: K-VRQOL의 표준화된 총점수; f) SE: social-emotional, 사회-심리영역 점수; g) PF: physical-functional, 신체기능 영역 점수

3. 음성장애군의 G 척도에 따른 K-VHI 및 K-VRQOL의 비교

음성장애군의 G 척도에 따른 K-VHI의 Total 점수는 통계적으로 유의미한 차이를 보였다($\chi^2 = 14.26$, $p = .003$). K-VHI의 세부영역별로 살펴보면, F 점수는 G 척도에 따라 유의미한 차이가 나타났으며($\chi^2 = 15.29$, $p = .002$), P 점수도 G 척도에 따라 유의미한 차이를 보였으며($\chi^2 = 16.66$, $p = .001$), E 점수 또한 G 척도에 따라 유의미한 차이가 있었다($\chi^2 = 9.55$, $p = .023$).

음성장애군의 G 척도에 따른 K-VRQOL의 Sum 점수도 통계적으로 유의미한 차이를 나타내었으며($\chi^2 = 15.03$, $p = .002$), 세부영역별로 SE 점수($\chi^2 = 11.18$, $p = .011$), PF 점수($\chi^2 = 15.26$, $p = .002$)도 G 척도에 따라 유의미한 차이를 보였다. 음성장애군의 G 척도에 따른 K-VHI와 K-VRQOL의 중위수와 사분위수는 <표 - 3>에 제시되었다.

<표 - 4> 음성장애군과 정상음성군의 K-VHI와 K-VRQOL

		음성장애군 총점수 (N = 120)	정상음성군 총점수 (N = 120)	t
K-VHI	Total ^{a)}	54.30 ± 27.32	17.58 ± 20.38	11.80**
	F ^{b)}	15.76 ± 9.69	10.34 ± 12.22	4.97**
	P ^{c)}	21.19 ± 9.09	4.97 ± 5.32	16.87**
	E ^{d)}	15.99 ± 11.78	2.26 ± 3.78	12.24**
Sum ^{e)}		63.73 ± 26.77	95.42 ± 7.95	-12.43**
K-VRQOL	SE ^{f)}	66.25 ± 29.27	96.93 ± 7.77	-11.10**
	PF ^{g)}	60.83 ± 26.15	94.41 ± 9.58	-13.21**

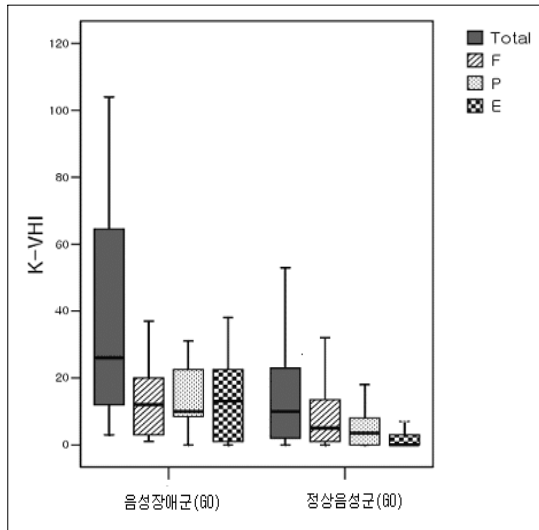
* $p < .05$, ** $p < .01$

a) Total: K-VHI의 총점수; b) F: functional, 기능 점수; c) P: physical, 신체 점수; d) E: emotional, 정서 점수; e) Sum: K-VRQOL의 표준화된 총점수; f) SE: social-emotional, 사회-심리영역 점수; g) PF: physical-functional, 신체기능 영역 점수

4. 음성장애군과 정상음성군의 K-VHI 및 K-VRQOL의 비교

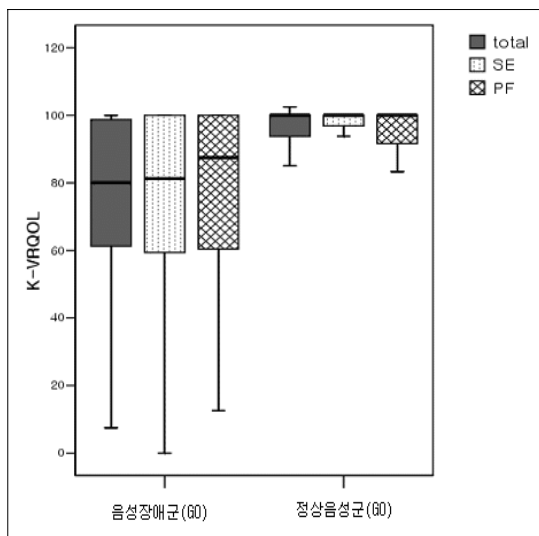
음성장애군(G0, G1, G2, G3)과 정상음성군(G0)의 K-VHI와 V-RQOL의 총점수 및 각 세부영역별 점

수는 모두 유의미한 차이를 보였으며(<표 - 4>), 이들 중 음성장애군의 G0 집단과 정상음성군(G0) 간의 K-VHI의 F 점수를 제외하고 Total 점수 ($Z = -2.50, p = .013$), P 점수 ($Z = -3.12, p = .002$) 그리고 E 점수 ($Z = -3.24, p = .001$)가 모두 유의미한 차이를 보였다(<그림 - 1>).



<그림 -1> 음성장애군 중 G0과 정상음성군(G0)의 K-VHI

음성장애군의 G0 집단과 정상음성군(G0)의 K-VRQOL의 Sum 점수 ($Z = -2.62, p = .009$), SE 점수 ($Z = -3.42, p = .001$) 그리고 PF 점수 ($Z = -2.22, p = .026$)가 모두 유의미한 차이를 보였다(<그림 - 2>).



<그림 - 2> 음성장애군 중 G0과 정상음성군(G0)의 K-VRQOL

5. 음성장애군의 진단별 분류에 따른 K-VHI 및 K-VRQOL 비교

진단별 분류에 따른 음성장애군의 K-VHI와 K-VRQOL의 평균과 표준편차는 <표 - 5>에 제시되었다. 이 중에서 G 척도를 통제한 후의 음성장애군의 진단별 분류에 따른 K-VHI 및 K-VRQOL의 총 점수와 세부영역별 점수의 차이를 살펴보면, K-VHI의 Total 점수는 통계적으로 유의미한 차이를 보였으며($F = 7.28, p = .001$), 사후검정에 의하면 양성 병변이 있는 구조적-기능적 장애 집단과 신경학적 장애 집단 간에 유의미한 차이가 있었으나($p = .001$), 양성 병변이 있는 구조적-기능적 장애 집단과 양성 병변이 없는 구조적-기능적 장애 집단 또는 양성 병변이 없는 구조적-기능적 장애 집단과 신경학적 장애 집단 간에는 유의미한 차이가 없었다. K-VHI의 세부 영역별 점수에서 F 점수($F = 10.42, p < .001$)와 E 점수($F = 7.85, p = .001$)는 음성장애군의 진단별 분류에 따라 유의미한 차이가 있었으나 P 점수는 유의미한 차이가 없었다. 사후검정에서 F 점수($p < .001$)와 E 점수($p < .001$)도 양성 병변이 있는 구조적-기능적 장애 집단과 신경학적 장애 집단 간에만 유의미한 차이를 나타내었다.

G 척도를 통제한 후의 음성장애군의 진단별 분류에 따른 K-VRQOL의 Sum 점수도 통계적으로 유의미한 차이를 보였다($F = 4.79, p = .010$). 진단별 분류에서 차이가 있는 집단들은 양성 병변이 있는 구조적-기능적 장애집단과 신경학적 장애 집단이었으며($p = .005$), 양성 병변이 있는 구조적-기능적 장애 집단과 양성 병변이 없는 구조적-기능적 장애 집단 또는 양성 병변이 없는 구조적-기능적 장애 집단과 신경학적 장애 집단 간에는 유의미한 차이가 없었다. K-VRQOL의 SE 점수($F = 8.06, p = .001$)와 PF 점수($F = 5.17, p = .007$)는 음성장애군의 진단별 분류에 따라 유의미한 차이가 있어 시행한 사후검정에 의하면, K-VRQOL의 모든 세부 영역별 점수에서 양성 병변이 있는 구조적-기능적 장애 집단과 신경학적 장애 집단에서만 통계적으로 유의미한 차이가 있었다 (SE 점수: $p < .001$; PF 점수: $p = .004$).

<표 - 5> 음성장애군의 진단별 분류에 따른 K-VHI 및 K-VRQOL 비교

		진단별 분류			F
		양성 병변이 있는 구조적-기능적 장애 집단 (N = 59)	양성 병변이 없는 구조적-기능적 장애 집단 (N = 29)	신경학적 장애 집단 (N = 32)	
K-VHI	Total ^{a)}	46.79 ± 25.11	53.13 ± 23.71	69.00 ± 29.16	7.28**
	F ^{b)}	14.41 ± 7.63	16.63 ± 7.84	22.66 ± 9.25	10.42**
	P ^{c)}	19.97 ± 9.67	20.27 ± 8.06	24.28 ± 8.42	2.24
	E ^{d)}	12.41 ± 9.88	16.23 ± 10.02	22.25 ± 13.69	7.82**
	Sum ^{e)}	70.47 ± 24.17	62.92 ± 24.00	52.27 ± 30.28	4.79*
K-VRQOL	SE ^{f)}	75.75 ± 25.62	54.17 ± 26.87	50.98 ± 31.52	8.06**
	PF ^{g)}	66.95 ± 24.69	62.08 ± 24.33	48.57 ± 26.91	5.17**

* $p < .05$, ** $p < .01$

a) Total: K-VHI의 총점수; b) F: functional, 기능 점수; c) P: physical, 신체 점수; d) E: emotional, 정서 점수; e) Sum: K-VRQOL의 표준화된 총점수; f) SE: social-emotional, 사회-심리영역 점수; g) PF: physical-functional, 신체기능 영역 점수

IV. 논의 및 결론

음성장애에 대한 객관적 평가와 평가자 혹은 대상자의 주관적 평가들 중 어느 한 가지만을 통해 복잡한 음성의 특성을 파악하기란 어렵다. 일반적으로 병원 환경에서 음성장애를 평가할 때 시간과 평가의 효율성을 위해 전문의나 음성언어치료사에 의해서만 음성을 평가하기 쉽다. 그러나 이들에 의한 음성평가는 대상자가 경험하는 음성과 관련된 일상생활에서의 신체적, 기능적 그리고 심리적인 불편감을 측정하기는 매우 어렵기 때문에 대상자 중심의 음성평가를 시행하는 것이 매우 중요하다. 대상자마다 음성 사용의 정도가 다르고, 음성을 전문적으로 사용하는지의 유무가 다르며, 음성 문제에 대한 인식이나 개인의 성격에 따라 자신의 음성장애의 정도를 다르게 평가하므로 대상자의 견해에서 본 음성평가를 시행함으로써 음성장애 치료의 궁극적 목표인 대상자의 음성에 대한 개별화된 만족도를 좀 더 정확하게 평가할 수 있을 것이다.

본 연구에 참여한 음성대상자의 진단별 분류를 살펴보면, 양성 병변이 있는 구조적-기능적 장애 집단은 58명, 양성 병변이 없는 구조적-기능적 장애 집단은 30명, 그리고 신경학적 장애 집단이 32명이었다. 각 진단별 분류에 따라 G 척도를 등급별로 나누어본 결과, G0과 G3 보다는 G1과 G2에 많은 대상자가 분포되어 있었다. 본 연구에서 이러한 대상자수의 불균형이 통계 결과에 영향을 미쳤을 수 있다는 제한점이 있다. G0에 해당하는 대상자가 적은 이유는 본 연구가 3차 기관인 종합병원에서 시행되었고, 음성정밀검

사를 시행하는 환자를 대상으로 이루어진 것으로, G0 집단은 자신의 음성에 민감하게 반응하거나 음성사용에 매우 중요성을 두고 있는 사람들로 구성되어 있기 때문에 많은 사람들이 이에 포함되기 어려웠을 것이다. 또한 G3 집단의 대상자가 적은 것은 평가자의 내적 규정과 관련이 있을 것이다(Hakkesteegt et al., 2008). 본 연구가 시행된 음성클리닉에는 매우 좋은 음성에서부터 음성이 거의 산출되지 않는 경우의 다양한 음성장애를 평가하므로 본 연구자의 엄격한 기준에 의해 매우 심각한 정도의 이상 음성 혹은 무성증의 경우에만 G3을 부여한 것과 관련이 있을 것이다.

선행연구들에 의하면 객관적 평가, 특히 음향학적 평가는 평가자의 청지각적 평가와 높은 상관관계를 보였으나 대상자의 주관적 평가와는 매우 낮은 상관성을 보였다(표화영 외, 1999; Bhuta, Patrick & Garnett, 2004; Hakkesteegt et al., 2008; Hsiung, Pai & Wang, 2002; Wheeler, Collins & Sapienza, 2005; Woisard et al., 2007). 이에 본 연구는 평가자의 청지각적 평가와 대상자의 주관적 평가 간의 상관성을 파악하고자 시행되었다. 이를 위해 평가자의 청지각적 평가로 GRBAS 척도 중 G 척도를 사용하였으며, 대상자의 주관적 평가로 설문지 형식의 K-VHI와 K-VRQOL이 사용되었다.

<표 - 1>에 의하면, G 척도와 K-VHI의 Total, F, P, E 점수와의 상관계수들은 모두 .320 이하로 약한 정적 상관관계를 나타냈으며, G 척도와 K-VRQOL의 Sum, SE, PF 점수와의 상관계수들도 .320보다 낮은 약한 부적 상관관계를 보였다. 이는 평가자 중심의

최도인 GRBAS와 대상자 중심의 척도인 V-RQOL 및 IPVI의 비교연구(Karnell et al., 2007)에서 제시된 낮은 상관관계($r = -.40 \sim -.57$)와 같은 결과를 보인다. 또한 객관적 평가와 청지각적 평가 간의 높은 상관관계를 보이는 선행연구와 객관적 평가와 대상자의 주관적 평가 간의 상관관계가 거의 없다는 연구 결과를 통해 추론된 가설과 일치한다. 이러한 결과는 평가자의 청지각적 평가와 대상자의 주관적 평가 사이의 상관관계가 매우 약함을 뜻하며, 다시 말해 평가자에 의한 청지각적 평가 점수가 높아도 주관적 평가 점수가 높다고 할 수 없다는 것이다. 따라서 대상자에 의해 산출된 음질이 대상자가 경험하는 음성장애 정도나 음성사용의 불편을 대변할 수 없으므로 음성을 총체적으로 평가하고자 한다면, 다양한 측면, 특히 대상자의 주관적 평가가 반드시 시행되어야 함을 시사한다.

더불어, GRBAS 척도 간의 상관관계를 살펴본 결과, G 척도는 R 척도($r = .453$) 및 B 척도($r = .584$)와 통계적으로 유의미한 중간 정도의 정적 상관관계를 보였으며, B 척도와 A 척도 또한 통계적으로 유의미한 중간 정도의 정적 상관관계($r = .564$)를 보였다. 이는 MDVP를 이용한 음향학적 평가와 GRBAS 척도 간의 상관관계를 조사한 표화영 외(1999)에서 G 척도와 R 척도 및 B 척도가 MDVP 상에서 공통된 변수들과 관계가 있어 GRB간의 상관성이 높음을 간접적으로 뒷받침하는 결과이다. 음질을 포괄적으로 평가하는 G 척도는 일반적으로 음성의 거친 정도와 기식 정도 및 그 외의 다른 음질의 특성을 함께 평가하는 척도이다. 따라서 G 척도와 R 척도 및 B 척도와 중간 이상의 높은 상관성을 보인다는 것은 그다지 놀랄 만한 결과는 아니다.

음성장애군의 G 척도에 따른 K-VHI 및 K-VRQOL의 총점수와 세부영역별 점수는 모두 통계적으로 유의미한 차이를 보였다. <표-3>을 보면, 청지각적 평가인 G 척도에서 음성장애의 중증도가 높을수록 K-VHI의 총점수와 세부영역별 점수가 높았으며, K-VRQOL의 총점수와 세부영역별 점수는 낮은 것을 볼 수 있었다. GRBAS 척도와 K-VHI나 K-VRQOL 간에 강한 상관관계는 보이지는 않지만, 전반적으로 청지각적 평가와 대상자들의 주관적 평가가 유사한 경향을 나타낸다는 것을 알 수 있다. 그렇다면 청지각적 평가와 주관적 평가 간의 낮은 상관관계를 초래한 원인은 무엇일까? 이는 G 척도에 따른 각 집단별 K-

VHI와 K-VRQOL의 총점수와 세부영역별 점수의 높은 표준편차들을 통해 알 수 있을 것이다. 이들의 평균 점수가 G 척도별로 다르다 하더라도 각 척도별 집단 내의 표준편차가 크다는 것은 모든 대상자들이 자신의 음질의 정도와 음성에 대한 만족도를 동일하게 평가하지는 않는다는 것이다. 이러한 결과를 종합해 볼 때, 청지각적으로 평가되는 음질이 나쁘다 하더라도 모든 대상자가 반드시 자신의 음성에 대해 부정적인 견해를 갖는다고 할 수는 없으나, 대체로 음질이 나쁠수록 자신의 음성에 대해 만족감을 느끼지 못한다는 것을 알 수 있다.

일반적으로 '정상 음성'이라는 개념은 매우 주관적이고 개인차가 클 수 있다. 음성장애군 중 G0 집단에서 평가자에 의한 청지각적 평가는 정상이지만 대상자 스스로가 음질이나 음성 사용에 있어 불편을 느끼는 대상자들로 구성되어 있으며, 이 집단과 정상음성군을 비교했을 때 K-VHI의 기능 점수를 제외한 K-VHI 및 K-VRQOL의 총점수와 세부영역별 점수들이 모두 유의미한 차이가 있었다(<그림-1>, <그림-2>). G0 집단의 음성장애군은 정상음성군에 비해 자신의 음질과 음성사용이 일상생활의 기능적 측면에는 영향을 미치지 않지만, 음질과 음성산출 과정에서 후두의 불편감이 발생하거나 음성 산출에 대한 만족도가 낮다는 것을 보여줌으로써, K-VHI와 K-VRQOL이 음성장애군 중 G0 집단이 경험하고 있는 신체적, 정신적 불편감을 적절하게 제시할 수 있는 문항들로 구성되어 있음을 알 수 있다.

<표-4>를 살펴보면, 정상음성군의 K-VHI 점수의 표준편차가 높는데 원점수를 살펴본 결과, 20대와 30대 초반 여성 6명의 K-VHI 점수가 60점 이상으로 높아 이 점수들이 높은 표준편차에 영향을 준 것으로 볼 수 있다. 이를 통해 정상 음성사용자라 하더라도 자신의 음성이나 음성사용에 있어 불편이나 만족감을 갖지 못하는 사람들이 있음을 알 수 있다. 따라서 이들에 대한 심리적 중재를 포함한 음성치료가 필요하며, 일반인들이 음성치료를 대해 좀 더 친근하게 접근할 수 있는 방법을 마련하는 것이 필요하다고 본다.

G 척도를 통제된 후에 실시한 음성장애군의 진단별 분류에 따른 차이를 살펴보면, 음성장애군에서 K-VHI의 P 점수를 제외한 K-VHI 모든 영역별 점수와 K-VRQOL의 모든 영역별 점수가 통계적으로 유의미한 차이를 보였다. 이는 진단명이 다르다 하더라도 대상자가 느끼는 음성과 관련된 신체적 불편감

의 차이는 크게 없으나 진단명에 따라 음성과 관련된 일상생활이나 타인과의 관계에서의 불편감과 이로 인해 음성장애의 정도를 다르게 인식하고 있음을 알 수 있다. 특히 신체 점수를 제외한 K-VHI의 다른 영역 모두에서 신경학적 장애 집단이 양성 병변이 있는 구조적-기능적 장애 집단에 비해 유의미하게 높은 점수를 보였고, K-VQOL의 모든 영역에서는 유의미하게 낮은 점수를 보였다. 그러나 양성 병변이 있는 구조적-기능적 장애 집단과 양성 병변이 없는 구조적-기능적 장애 집단 또는 양성 병변이 없는 구조적-기능적 장애 집단과 신경학적 장애 집단 사이에는 K-VHI 및 K-VRQOL의 모든 영역에서 통계적으로 유의미한 차이가 없었다. 이를 살펴볼 때, 본 연구에서 분류된 양성 병변이 있는 구조적-기능적 장애는 음성치로나 성대미세수술 등을 통해 음성이 호전될 수 있는 가능성이 많은 음성장애로, 이를 경험하는 대상자는 장애 특성상 신경학적 장애를 경험하는 대상자들에 비해 자신의 음성 문제를 크게 염려하거나 불편해 하지 않는다고 판단할 수 있다. 이에 반해 성대마비나 경련성 발성장애와 같은 신경학적 장애 집단의 대상자들은 자신들의 음성장애를 완치하지 못하고 영구적으로 지닐 수도 있을 것이라는 부정적인 견해를 더 많이 지니고 있음을 알 수 있다.

본 연구는 평가자의 청지각적 평가와 대상자의 주관적 평가 간의 상관성을 분석하고, 청지각적 평가 중의 하나인 GRBAS 척도 중 음질을 종합적으로 평가할 수 있는 G 척도의 등급에 따라 주관적 평가의 차이가 있는지를 살펴보았다. 결론적으로, 청지각적 평가와 주관적 평가를 비교하였을 때, 청지각적 평가에서 음성장애의 정도가 높을 때 주관적으로 인지하는 음성장애의 정도도 높다고 말할 수는 있으나, 이들 간에 상관관계는 낮아 평가자가 평가한 음성장애의 정도와 대상자가 지각하고 있는 음성장애의 정도가 항상 일치한다고는 볼 수 없다. 또한 진단별 분류에 따른 대상자들의 주관적 평가 또한 많은 차이가 있었다. 전반적으로 구조적-기능적 음성장애를 지닌 대상자에 비해 신경학적 장애를 지닌 대상자들이 자신들의 음성장애에 대한 불만과 일상생활에서의 불편감을 많이 지각하고 있음을 알 수 있었다. 따라서 음성치로나 기타 치료를 시행할 때 이들의 음성장애에 대한 주관적 평가를 반영함으로써 치료에 대한 효과를 증대시킬 수 있어야 할 것이다. 더불어 음성은 음도, 음의 강도와 음질로 구성되는 다차원적인 성질을 지니기 때문

에 하나의 음성평가로는 정확하게 진단할 수 없으며 적합한 치료를 시행할 수 없으므로 음성평가를 시행할 때 객관적 음성평가만 시행할 것이 아니라 평가자의 청지각적 평가 및 대상자의 주관적인 음성평가를 반드시 시행하는 다양한 측면의 음성평가가 요구되며, 추후 연구로써 청지각적 평가와 주관적 평가뿐만 아니라 이들과 객관적 평가 간의 종합적인 상관성을 연구함으로써 보다 정확하게 음성장애를 평가하는 방법을 간구해 볼 필요가 있을 것이다.

참 고 문 헌

- 김재옥·임성은·박선영·최성희·최재남·최홍식(2007a). 한국어판 음성장애지수와 음성관련 삶의 질의 타당도 및 신뢰도 연구. 『음성과학』, 14(3), 111-125.
- 김재옥·최성희·임성은·최재남·최홍식(2007b). 전문직 음성사용자의 주관적 음성평가도구간의 비교. 『대한음성언어의학회 학술대회 발표논문집』. 그랜드 힐튼호텔, 서울.
- 박영학·이정학·주영훈·박성신·방충일·김민식·조승호(2005). 양성 성대 질환 환자의 후두 미세 수술전후 음성 장애 지수 및 음성 분석의 유용성. 『대한음성언어의학회지』, 16(1), 23-27.
- 엄재옥·한태희·이종락·김영완(2007). 양성 후두 병변에 대한 Accent Method의 음성치료 효과. *Journal of Clinical Otolaryngology*, 18(2), 195-201.
- 표화영·최성희·임성은·심현섭·최홍식·김광문(1999). 성대 폴립 환자를 대상으로 한 GRBAS 척도와 MDVP 측정치 간의 상관관계 연구. 『대한음성언어의학회지』, 10(2), 154-163.
- American Speech-Language-Hearing Association (2002). *Consensus Auditory Perceptual Evaluation of Voice (CAPE-V)*. Pittsburgh, PA: Author.
- Anders, L. C., Hollien, H., Hurme, D., Sonnine, A., & Wendler, J. (1988). Perception of hoarseness by several classes of listeners. *Folia Phoniatria*, 40(2), 91-100.
- Bhuta, C., Patrick, L., & Garnett, J. (2004). Perceptual evaluation of voice quality and correlation with acoustic measurements. *Journal of Voice*, 18(3), 299-304.
- Colton, R. H., & Casper, J. K. (1996). *Understanding voice problems: A physiological perspectives for diagnosis and treatment* (2nd ed.). Philadelphia, PA: Lippincott Williams & Wilkins.
- De Bodt, M. S., Wuyts, F. L., Van de Heyning, P. H., & Croux, C. (1997). Test-retest study of GRBAS scale: Influence of experience and professional background on perceptual rating of voice quality. *Journal of Voice*, 11(1), 74-80.
- Deary, I. I., Wilson, J. A., Carding, P. N., & MacKenzie, K.

- (2003). VoiSS: A patient-derived Voice Symptom Scale. *Journal of Psychosomatic Research*, 54, 483-489.
- Eom, J. W., Han, T. H., Lee, J. B., & Kim, Y. W. (2007). Efficacy of the Accent Method of voice therapy on benign lesions of vocal cord. *Journal of Clinical Otolaryngology*, 18(2), 195-201.
- Hakkesteegt, M. M., Brocaar, M. P., Wieringa, M. H., & Feenstra, L. (2008). The relationship between perceptual evaluation and objective multiparametric evaluation of dysphonia severity. *Journal of Voice*, 22(2), 138-145.
- Hirano, M. (1981). *Clinical examination of voice*. New York: Springer-Verlag.
- Hogikyan, N. D., & Sethuraman, G. (1999). Validation of an instrument to measure Voice-Related Quality of Life (V-RQOL). *Journal of Voice*, 13(4), 557-569.
- Hsiung, M. W., Pai, L., & Wang, H. W. (2002). Correlation between voice handicap index and voice laboratory measurements in dysphonic patients. *European Archives of Oto-Rhino-Laryngology*, 259(2), 97-99.
- Jacobson, B. H., Johnson, A., Grywalski, C., Silbergleit, A., Jacobson, G., Benninger, M. S., & Newman, C. W. (1997). The Voice Handicap Index (VHI): Development and validation. *American Journal of Speech-Language Pathology*, 6, 66-70.
- Karnell, M. P., Melton, S. D., Childes, J. M., Coleman, T. C., Dailey, S. A., & Hoffman, H. T. (2007). Reliability of clinician-based (GRBAS and CAPE-V) and patient-based (V-RQOL and IPVI) documentation of voice disorders. *Journal of Voice*, 21(5), 576-590.
- Kreiman, J., Gerratt, B. R., Kempster, G. B., Erman, A., & Berke, G. S. (1993). Perceptual evaluation of voice quality: Review, tutorial, and a framework for future research. *Journal of Speech and Hearing Research*, 36(1), 21-40.
- Ma, E., & Yiu, E. (2006). Multiparametric evaluation of dysphonic severity. *Journal of Voice*, 20(3), 380-390.
- Thomas, G., Kooijman, P. G., Donders, A. R., Cremers, W. R., & de Jong, F. I. (2007). The voice handicap of student-teachers and risk factors perceived to have a negative influence on the voice. *Journal of Voice*, 21(3), 325-336.
- Wheeler, K., Collins, S. P., & Sapienza, C. M. (2005). The relationship between VHI scores and specific acoustic measures of mildly disordered voice production. *Journal of Voice*, 20(2), 308-317.
- Wingate, J. M., Brown, W. S., Shrivastav, R., Davenport, P., & Sapienza, C. M. (2007). Treatment outcomes for professional voice users. *Journal of Voice*, 21(4), 433-449.
- Woisard, V., Bodin, S., Yardeni, E., & Puech, M. (2007). The voice handicap index: Correlation between subjective patient response and quantitative assessment of voice. *Journal of Voice*, 21(5), 623-631.

ABSTRACT

Comparison of Clinicians' Perceptual Evaluations and Patients' Subjective Evaluations of Voice Disorders

Jaeock Kim^{a,§} · Hong-Shik Choi^b

^a Major in Speech Pathology Education, Graduate School of Education, Kangnam University, Yongin, Korea

^b Department of Otorhinolaryngology, The Institute of Logopedics & Phoniatrics, Yonsei University College of Medicine, Seoul, Korea

Background & Objectives: Among the different types of voice evaluations, objective quantitative evaluations and clinicians' perceptual evaluations are commonly used. However, patients' subjective evaluations must be considered to accomplish a more accurate and total assessment of voice disorders. The purpose of this study was to examine the relationship between clinicians' perceptual evaluations and patients' subjective evaluations of voice disorders. **Methods:** We enrolled 120 people with voice disorders and 120 people without voice disorders and asked them to fill out Korean Voice Handicap Index (K-VHI) and Korean Voice Related Quality of Life (K-VRQOL) forms. The patients' voices were perceptually evaluated by an experienced speech-language pathologist and an otolaryngologist using GRBAS scales. **Results:** There was a low degree of correlation between the G scale on the GRBAS scales and the K-VHI and K-VRQOL self-reports. However, mean scores on both total and subscales in K-VHI and K-VRQOL decreased as the severity of voice disorders on the G scale increased. In particular, G3 differed significantly from G0, G1, and G2 in the results of both K-VHI and K-VRQOL. In addition, patients who were diagnosed differently had significantly different total and subscale scores on the K-VHI and K-VRQOL, with the exception of the physical score on the K-VHI. **Discussion & Conclusion:** There was no significant relationship between clinicians' perceptual evaluations and patients' subjective evaluations of voice disorders. The severity of voice disorders assessed by clinicians differed from the voice handicap and voice-related quality of life perceptions of patients. Therefore, there is a need to use various assessment methods when evaluating voice disorders, including patients' subjective evaluations. (*Korean Journal of Communication Disorders* 2009;14;223-235)

Key Words: voice disorders, perceptual voice evaluation, patient's subjective evaluation, GRBAS

[§] Correspondence to

Prof. Jaeock Kim, PhD,
Major in Speech Pathology
Education, Graduate School of
Education, Building, Kangnam
University, 111 Gugal-Dong,
Kihung-Gu, Yongin-Si,
Kyungki-Do, Korea
e-mail: jaeock@gmail.com
tel.: +82 31 280 3221

REFERENCES

- American Speech-Language-Hearing Association (2002). *Consensus Auditory Perceptual Evaluation of Voice (CAPE-V)*. Pittsburgh, PA: Author.
- Anders, L. C., Hollien, H., Hurme, D., Sonnine, A., & Wendler, J. (1988). Perception of hoarseness by several classes of listeners. *Folia Phoniatrica*, 40(2), 91-100.
- Bhuta, C., Patrick, L., & Garnett, J. (2004). Perceptual evaluation of voice quality and correlation with acoustic measurements. *Journal of Voice*, 18(3), 299-304.
- Colton, R. H., & Casper, J. K. (1996). *Understanding voice problems: A physiological perspectives for diagnosis and treatment* (2nd ed.). Philadelphia, PA: Lippincott Williams & Wilkins.
- De Bodt, M. S., Wuyts, F. L., Van de Heyning, P. H., & Croux, C. (1997). Test-retest study of GRBAS scale: Influence of experience and professional background on perceptual rating of voice quality. *Journal of Voice*, 11(1), 74-80.

* This work was supported by the Korea Research Foundation Grant funded by the Korean Government (MOEHRD) (KRF-2007-355-E00012).

■ Received October 21, 2008 ■ Final revision received June 15, 2009 ■ Accepted June 17, 2009.

© 2009 The Korean Academy of Speech-Language Pathology and Audiology <http://www.kasa1986.or.kr>

- Deary, I. I., Wilson, J. A., Carding, P. N., & MacKenzie, K. (2003). VoiSS: A patient-derived voice symptom scale. *Journal of Psychosomatic Research, 54*, 483-489.
- Eom, J. W., Han, T. H., Lee, J. B., & Kim, Y. W. (2007). Efficacy of the Accent Method of voice therapy on benign lesions of vocal cord. *Journal of Clinical Otolaryngology, 18*(2), 195-201.
- Hakkesteegt, M. M., Brocaar, M. P., Wieringa, M. H., & Feenstra, L. (2008). The relationship between perceptual evaluation and objective multiparametric evaluation of dysphonia severity. *Journal of Voice, 22*(2), 138-145.
- Hirano, M. (1981). *Clinical examination of voice*. New York: Springer-Verlag.
- Hogikyan, N. D., & Sethuraman, G. (1999). Validation of an instrument to measure Voice-Related Quality of Life (V-RQOL). *Journal of Voice, 13*(4), 557-569.
- Hsiung, M. W., Pai, L., & Wang, H. W. (2002). Correlation between voice handicap index and voice laboratory measurements in dysphonic patients. *European Archives of Oto-Rhino-Laryngology, 259*(2), 97-99.
- Jacobson, B. H., Johnson, A., Grywalski, C., Silbergleit, A., Jacobson, G., Benninger, M. S., & Newman, C. W. (1997). The Voice Handicap Index (VHI): Development and validation. *American Journal of Speech-Language Pathology, 6*, 66-70.
- Karnell, M. P., Melton, S. D., Childes, J. M., Coleman, T. C., Dailey, S. A., & Hoffman, H. T. (2007). Reliability of clinician-based (GRBAS and CAPE-V) and patient-based (V-RQOL and IPVI) documentation of voice disorders. *Journal of Voice, 21*(5), 576-590.
- Kim, J., Choi, S. H., Lim, S. E., Choi, J. N., Choi, H. S. (2007b). Comparison among the self-assessment voice rating scales in professional voice users. *Proceedings of the Korean Society of Otorhinolaryngology-Head and Neck Surgery 81th Annual congress*. Grand-Hilton Hotel, Seoul.
- Kim, J., Lim, S. E., Park, S. Y., Choi, S. H., Choi, J. N., Choi, H. S. (2007a). Validity and reliability of Korean-version of Voice Handicap Index and Voice-Related Quality of Life. *Speech Sciences, 14*(3), 111-125.
- Kreiman, J., Gerratt, B. R., Kempster, G. B., Erman, A., & Berke, G. S. (1993). Perceptual evaluation of voice quality: Review, tutorial, and a framework for future research. *Journal of Speech and Hearing Research, 36*(1), 21-40.
- Ma, E., & Yiu, E. (2006). Multiparametric evaluation of dysphonic severity. *Journal of Voice, 20*(3), 380-390.
- Park, Y. H., Lee, J. H., Joo, Y. H., Park, S. S., Bang, C. I., Kim, M. S., Cho, S. H. (2005). Validity of Voice Handicap Index and voice analysis following laryngeal micro-surgery for benign vocal cord lesions. *The Journal of the Korean Society of Logopedics and Phoniatics, 16*(1), 23-27.
- Pyo, H. Y., Choi, S. H., Lim, S. E., Sim, H. S., Choi, H. S., Kim, K. M. (1999). The correlation between GRBAS scales and MDVP parameters on the pathologic voices of the patients with vocal polyps. *The Journal of the Korean Society of Logopedics and Phoniatics, 10*(2), 154-163.
- Thomas, G., Kooijman, P. G., Donders, A. R., Cremers, W. R., & de Jong, F. I. (2007). The voice handicap of student-teachers and risk factors perceived to have a negative influence on the voice. *Journal of Voice, 21*(3), 325-336.
- Wheeler, K., Collins, S. P., & Sapienza, C. M. (2005). The relationship between VHI scores and specific acoustic measures of mildly disordered voice production. *Journal of Voice, 20*(2), 308-317.
- Wingate, J. M., Brown, W. S., Shrivastav, R., Davenport, P., & Sapienza, C. M. (2007). Treatment outcomes for professional voice users. *Journal of Voice, 21*(4), 433-449.
- Woisard, V., Bodin, S., Yardeni, E., & Puech, M. (2007). The voice handicap index: Correlation between subjective patient response and quantitative assessment of voice. *Journal of Voice, 21*(5), 623-631.