

한국어판 음성장애지수
(Korean-Voice Handicap Index):
번안본 검증 및 새 지수 개발

연세대학교 대학원
언어병리학 협동과정
윤 영 선

한국어판 음성장애지수
(Korean-Voice Handicap Index):
번안본 검증 및 새 지수 개발

지도 김 향 희 교수

이 논문을 박사 학위논문으로 제출함

2007년 6월 일

연세대학교 대학원

언어병리학 협동과정

윤 영 선

윤영선의 박사 학위논문을 인준함

심사위원 김 향 희 인

심사위원 최 홍 식 인

심사위원 손 영 익 인

심사위원 박 은 숙 인

심사위원 신 지 철 인

연세대학교 대학원

2007년 6월 일

감사의 글

연세 동산에서 언어치료교육과정을 거쳐 학생 신분의 마지막 과정인 박사학위 과정까지 마치게 하신 하나님의 은혜에 먼저 감사드립니다. 어언 십 년이 넘는 시간 동안 학업과 언어치료사의 일을 감당하면서, 부족한 논문이지만 이것으로 하나의 결실을 얻게 되어 감회가 새롭습니다.

논문이 나오기까지 지도해주신 심사위원 교수님들께 진심으로 감사드립니다. 병원 언어치료사의 역할을 처음 가르쳐 주셨고 그 인연이 계속되어 오늘날 지도 교수까지 담당해주신 김향희 교수님, 수업 때 음성장애 학문을 처음 접하게 하시고 지금껏 그 분야의 대가로서 후학들을 격려해주시는 최홍식 교수님, 지난 십여 년 간 임상 현장에 함께 일을 하면서 진정한 음성언어치료사로 만들어주신 손영익 교수님, 연구자의 시각으로 논문에 대한 명쾌한 고견을 주신 박은숙 교수님, 논문의 구석구석을 꼼꼼히 살펴주시고 세심한 조언을 해주신 신지철 교수님께 다시 한번 머리 숙여 고마운 맘을 전해드립니다. 또한 논문 자료를 모을 수 있도록 도와주신 각 병원 기관과 참여하신 언어치료사 선생님들, 그리고 설문 대상이 되어주신 분들께 지면을 통해 깊이 감사드립니다. 일일이 다 적지는 못하지만, 제가 논문을 잘 마칠 수 있도록 후원해준 자랑스럽고 든든한 우리 세언회 동료 및 후배들께도 특별히 감사드립니다.

더불어 기쁨을 함께 나누고 싶은 분들이 있습니다. 제게 늘 관심을 두고 상담 해주시는 장병기 교수님, 심현섭 교수님 그리고 가르침을 주신 학창시절의 은사님들, 격려해주신 삼성서울병원 이비인후과 교수님들과 동료직원들, 늘 기도해주시는 아현감리교회 목사님과 믿음의 식구들, 시온찬양대, 친애하는 벗과 지인들 모두께 감사드립니다. 끝으로, 사랑으로 저를 지켜봐주시고 함께 기뻐해주신 부모님, 저의 학위취득 날을 기대하며 연명하시는 조부모님, 연구에 많은 도움을 준 오빠 윤동필 교수를 포함하여 멀리 미국에서 기도해준 저의 모든 가족 일원께.....사랑하고 그리고 감사합니다.

저자 씀

차 례

국문 요약	1
I. 서 론	4
1. 음성장애관련 심리측정적 평가법의 필요성	4
2. VHI 및 기타 평가도구	6
3. 공인타당도 검증 문제	9
4. 질환군별 VHI 결과 비교	10
5. VHI와 청지각적/음향학적 검사 결과 간의 비교	11
6. VHI-10의 유용성	12
7. VHI 번안본의 검증 필요성	12
8. 문화적 차이를 반영하는 문항의 필요성	13
9. 연구의 목적	14
II. 대상 및 방법	17
1. 연구 대상	17
가. 환자실험군	17
나. 정상대조군	18
2. 연구 방법	20
가. 한국어판 VHI 제작	20
나. 설문지 구성	20
다. 설문지 시행 및 수집	21
(1) VHI, IPVI 및 SF-36	21
(2) 검사-재검사	21
(3) VHI-10 시행	22
라. 음향학적 검사	22
3. 통계 분석	22
III. 연구 결과	23
1. VHI 신뢰도 및 타당도	23
가. VHI 결과 분석	23

나. 신뢰도	27
(1) 검사-재검사 신뢰도	27
(2) 문항내적일관성	27
(3) 문항변별도	28
다. 타당도	30
(1) 음성장애군과 정상군의 변별성	30
(2) 공인타당도	31
(3) 요인분석	32
2. VHI-30과 SF-36의 상관성	34
가. SF-36 결과분석	34
나. 상관 분석	38
(1) VHI와 SF-36 간의 총점 비교	38
(2) VHI 및 SF-36 간의 하위 항목 비교	39
3. 질환별 VHI 및 SF-36 결과 비교	40
가. 질환 분류	40
나. 질환별 VHI 결과 비교	40
다. 질환별 SF-36 결과 비교	40
4. VHI-30과 음향학적 결과 비교	43
가. 음향학적 결과 분석	43
나. VHI-30과 음향학적 결과 간의 상관성	43
5. VHI-30과 VHI-10 비교	44
가. VHI-10 신뢰도	44
(1) VHI-30에서 추출한 VHI-10과 VHI-10 간의 일치도	44
(2) 문항내적일관성	45
나. VHI-10 타당도	45
(1) 환자군과 정상군의 총점 비교	46
(2) 문항변별도	46
(3) 공인타당도	47
(4) 요인분석	48
(5) 문항 간 상관분석	48
다. VHI-30과 VHI-10 상관성	49
라. VHI-30에 대한 총점 비율	50
6. 타당성이 큰 10문항 선별	50
가. 문항 분석	50

(1) 문항별 환자군-정상군 간 평균 차 비교	50
(2) 문항변별도	52
(3) 문항별 Cronbach's α	52
나. 새로 선택된 10문항(nVHI-10)	54
(1) VHI-30과 nVHI-10의 상관성	54
(2) VHI-30에 대한 총점 비율	55
(3) 공인타당도	55
(4) 요인분석	55
(5) 문항 간 상관분석	55
7. '노래 부르기' 문항 분석	57
가. 문항 분석	57
(1) 문항별 환자군-정상군 간 평균 차 비교	57
(2) 문항변별도	59
(3) 문항별 Cronbach's α	59
나. '노래 부르기'가 포함된 10문항(sVHI-10) 제작	61
(1) VHI-30에 대한 총점 비율	61
(2) 공인타당도	61
(3) 요인분석	61
(4) 문항 간 상관분석	62
IV. 고 찰	64
한국어판 VHI의 유용성	64
문항수를 단축한 VHI	65
본 연구의 제한점 및 향후 과제	68
V. 결 론	70
참고 문헌	71
부록(설문지 예)	77
영문 요약	86

그림 차례

그림 1. 환자군의 VHI-30의 총점 분포	25
그림 2. 정상군의 VHI-30의 총점 분포	26
그림 3. VHI-30의 검사-재검사 일치도	27
그림 4. 환자군과 정상군의 VHI-30 총점 비교	30
그림 5. 환자군 및 정상군의 VHI-30 총점과 IPVI 1번 문항 점수 간의 상관성	31
그림 6. VHI-30의 요인분석에 따른 scree plot	32
그림 7. 환자군의 SF-36 총점 분포	35
그림 8. 정상군의 SF-36 총점 분포	36
그림 9. 환자군과 정상군의 SF-36 총점 비교	37
그림 10. VHI-30과 SF-36의 총점 간 상관성	38
그림 11. 질환별 환자군과 정상군의 VHI-30 총점 비교	41
그림 12. 질환별 환자군과 정상군의 SF-36 총점 비교	42
그림 13. VHI-10의 일차 시행과 VHI-10 이차 시행 간의 일치도	45
그림 14. 환자군과 정상군 간의 VHI-10 총점 비교	46
그림 15. VHI-10 총점과 IPVI 1번 문항 점수 간의 상관성	48
그림 16. VHI-30 총점과 VHI-10 총점 간의 상관성	49
그림 17. VHI-30 총점과 nVHI-10 총점 간의 상관성	54
그림 18. nVHI-10 총점과 IPVI 1번 문항 점수 간의 상관성	56
그림 19. sVHI-10 총점과 IPVI 1번 문항 점수 간의 상관성	62

표 차례

표 1. 환자군과 정상군의 연령 및 성별 분포	19
표 2. 환자군의 음성장애별 분포	19
표 3. 환자군과 정상군의 VHI-30 문항별 평균과 표준편차	24
표 4. VHI-30의 각 문항과 총점 간의 상관계수	29
표 5. VHI-30의 요인분석 결과표	33
표 6. VHI-30 및 SF-36의 하위 항목 간 상관계수	39
표 7. 질환별 환자군과 정상군의 VHI-30 총점의 평균	41
표 8. 질환별 환자군과 정상군의 SF-36 총점의 평균	42
표 9. 환자군의 MDVP 파라미터 값	43
표 10. VHI-30 총점과 음향학적 파라미터 간의 상관계수	44
표 11. VHI-10 총점과 각 문항 간 상관계수	47
표 12. VHI-10 문항 간 상관계수	49
표 13. VHI-30에 대한 환자군-정상군의 평균 차	51
표 14. VHI-30에 대한 문항-총점 상관성과 삭제 시 Cronbach's α	53
표 15. nVHI-10 문항 간 상관계수	56
표 16. VHI-30 및 '노래 부르기' 세 문항에 대한 환자군-정상군의 평균 차	58
표 17. VHI-30 및 '노래 부르기' 세 문항에 대한 문항-총점 상관성과 삭제 시 Cronbach's α	60
표 18. sVHI-10 문항 간 상관계수	63

부록(설문지 예)

부록 1. IPVI(Iowa Patient's Voice Index) 중 1번 문항	77
부록 2. VHI(Voice Handicap Index)-30	78
부록 3. '노래 부르기' 문항	79
부록 4. SF-36	80
부록 5. VHI(Voice Handicap Index)-10	84
부록 6. 새로 선택한 VHI-10 문항(nVHI-10)	85
부록 7. '노래 부르기'가 포함된 VHI 10 문항(sVHI-10)	85

한국어판 음성장애지수(Korean-Voice Handicap Index): 번안본 검증 및 새 지수 개발

음성 문제를 평가하고 치료의 호전도를 알아보기 위해서는 평가자의 주관적 및 객관적인 평가뿐만 아니라 환자의 주관적인 평가가 병행되어야 한다. 우리나라에서는 최근에서야 음성장애에 관련된 심리측정적 평가(voice-related psychometric evaluation)도구를 사용하기 시작하였으나, 신뢰도 및 타당도에 대한 검증이 아직 이루어지지 않았다.

VHI(Voice Handicap Index)는 표준화된 도구는 아니지만, 음성장애환자에서 삶의 질과 연관된 주관적인 평가로서 가장 많이 사용되어온 평가도구이다. 본 연구에서는 후두암을 제외한 음성장애환자 156명과 정상군 66명을 대상으로 하여, 한국어판 VHI 30문항(VHI-30) 및 VHI-10의 신뢰도와 타당도를 제시하였다. 또한, 30 문항 중 새로 10문항(nVHI-10)을 추려보고, 노래 부르기 3문항을 첨부하여 총 33문항 중 10문항(sVHI-10)을 추려보아, 언어 및 문화적 차이로 인한 설문 결과의 차이가 있는지 알아봄과 동시에, 설문에 가장 효과적인 문항 구성을 제시해보고자했다. 시행된 설문은 VHI-30, 한국의 문화적 특성을 반영한 ‘노래 부르기’ 세 문항, IPVI(Iowa Patient’s Voice Index) 1번 문항, SF(Short Form)-36이었다. 결과는 다음과 같았다.

첫째, 내용타당도를 거친 VHI-30의 번역본에 대한 신뢰도는 피험자 내 일치도가 높았고($r=.971, p<0.01$), 문항내적일관성도 매우 높았다(Cronbach’s $\alpha=0.969$). 또한 타당도에서는 환자군이 정상군에 비해 유의하게 총점이 높고, 문항별 점수와 총점 간의 상관성이 높았으므로($r=0.583\sim 0.846, p<0.01$), 음성 문제에 대한 변별력이 높은 검사 도구임을 보였다. 공인타당도를 대신하여 IPVI의 1번 문항($r=0.664, p<0.01$)과 SF-36 간의 상관성($r=0.372, p<0.01$)은 모두 선행 연구와 같은 정도의

상관성을 보였다. 또한, 요인분석 상 세 가지의 하위항목으로 구분이 가능하였다.

둘째, 음성장애 질환군에 따라 VHI-30 총점에 유의한 차이를 보여, 집단 간 비교가 가능함을 보였다. 질환군은 신경학적 장애, 기능적 장애, 구조적 장애, 염증성 장애로 나누어 비교하였고, 선행연구와 마찬가지로 신경학적 장애는 기타 장애군보다 음성장애로 인한 삶의 질이 가장 낮은 집단임을 나타내었다.

셋째, VHI-30과 MDVP(Multi-Dimensional Voice Program) 파라미터들 간의 상관성은 없거나 낮았으므로, 환자의 주관적 음성 문제 인식 정도와 객관적 평가에 의한 음성 문제 정도는 그 기준에 차이가 있음을 보였다.

넷째, VHI-10 또한 VHI-30과 마찬가지로 검증을 통해 신뢰도 및 타당도가 있는 평가도구임을 확인하였다. VHI-30과 VHI-10 간의 상관성은 매우 높았다 ($r=0.968$, $p<0.01$). VHI-10 평균 총점 값 대 VHI-30 평균 총점 값 비율이 0.333보다 높은 0.368이었다. VHI-10은 요인분석 상 단일항목을 보였다.

다섯째, VHI-30 문항에서 타당성이 높은 10개의 문항을 새로 추려본 결과, nVHI-10에 해당하는 것은 P13, E23, P10, P14, P20, E25, E24, P26, E7, F16이었다. 이 중 기존 VHI-10과 중복이 되는 문항은 5개이었다. nVHI-10 평균 총점 값 대 VHI-30 평균 총점 값 비율이 0.408로, 기존 VHI-10의 비율(0.368)보다 높았다.

여섯째, VHI-30 문항과 한국의 문화적 특성을 반영한 ‘노래 부르기’의 세 문항을 포함하여 문항 분석을 한 경우에, sVHI-10에 해당하는 것은 P13, E23, P17, P10, P14, P20, E25, E24와 신체적인 면의 S31(‘노래 부르기가 힘들다’) 및 기능적인 면의 S32(‘사람들과 어울릴 때 노래를 잘 부르지 못해 불편할 때가 있다’)이었다. sVHI-10 평균 총점 값 대 VHI-30 평균 총점 값 비율이 0.456으로, 기존 VHI-10의 비율(0.368) 및 nVHI-10의 비율(0.408)보다 높았다.

이상의 결과로, 한국어판 VHI-30 및 VHI-10은 음성장애 심리측정적 평가도구로서 신뢰도 및 타당도가 있음을 밝혔다. 기존의 VHI-10에 비하여 본 연구에서 제시한 nVHI-10과 sVHI가 상대적으로 민감도가 더 높았으므로, 언어 및 문화권의 차이로

인한 문항 선정 기준에 차이가 있음을 제시해보았다. 따라서 향후 음성장애에 관련된 문항 제작을 하거나, 결과를 해석할 경우에는 언어 및 문화권의 차이에 대한 고려를 하는 것이 바람직하리라 생각한다.

핵심되는 말: 음성장애(voice disorders), 음성장애지수(VHI: Voice Handicap Index), 삶의 질(quality of life), SF-36, 교차문화적 연구(cross-cultural research)

한국어판 음성장애지수 (Korean-Voice Handicap Index): 번안본 검증 및 새 지수 개발

<지도교수 김 향 희>

연세대학교 대학원 언어병리학협동과정
윤 영 선

I. 서 론

1. 음성장애관련 심리측정적 평가법의 필요성

음성장애(voice disorders)¹⁾는 호흡, 발성, 그리고 공명의 과정에서 성도(vocal tract)의 어느 한 곳이라도 구조적 또는 기능적 이상이 나타난 결과로서, 화자의 나이, 성별, 지리적 배경에 기준했을 때 강도(loudness), 음도(pitch), 음질(quality) 및 공명(resonance) 중 하나 또는 그 이상이 정상 범주에서 벗어난 것을 말한다.¹⁾ 의사소통 기능에 기준해서는 너무 크거나 작은 목소리, 성대를 상하게 할 수 있는 목소리, 듣기에 거슬리는 목소리, 음도를 다양하게 구사하지 못하여 감정이나 뜻이 전달되지 못하는 목소리, 청자가 화자의 나이와 성별을 잘못 판단하게 하는 목소리 등도 음성장애라고 지칭할 수 있다.¹⁾

1) 논문 제목에 한하여 음성장애지수(Voice Handicap Index)의 'handicap'을 '장애'로 번역하였고, 본문에서는 '헨디캡'으로 표기함. 음성장애(voice disorders)의 'disorders'는 '장애'로 번역함.

일반적으로 음성장애를 평가하는 방법은 청지각적인 평가(auditory-perceptual evaluation),¹⁻³ 음향학적 평가(acoustic analysis),^{1,2,4} 공기역학적 평가(aerodynamic study),^{1,2} 생리학적인 평가인 후두스트로보스코피(laryngeal stroboscopy)^{1,2,5} 등으로, 평가자의 주관적 및 객관적 기준에 의해 시행된다. 임상에서 주로 사용하고 있는 청지각적 평가는 일본 음성언어의학회 위원회에서 제시한 GRBAS⁶로서 전체적인 중증도 G(grade)와 4가지의 대표적인 음질인 R(rough, 거친 음성), B(breathy, 바람 새는 음성), A(asthenic, 약한 음성), S(strained, 긴장하는 음성)을 기준하여 훈련된 검사자의 귀를 통해 정도를 판별하여 4점 척도로 측정한다. 음향학적인 평가는 MDVP(Multi-Dimensional Voice Program, Kay Elemetrics Corp., USA)와 같은 음향분석 프로그램을 이용하여 성대 진동에 따른 주파수 및 진폭의 안정성이나 잡음의 정도 등을 알아보는 방법으로, jitter⁴, shimmer⁴, NHR(noise to harmonic ratio) 등의 파라미터 수치가 정상 기준치에 비해 어느 정도 증가했는지를 알아본다. MDVP에서는 jitter의 기준치는 1.04%이고, shimmer의 기준치는 3.81%, NHR의 기준치는 0.19이며, 이들 기준치보다 수치가 상대적으로 증가할수록 음성에 문제가 있다고 해석한다. 공기역학적 검사는 초시계 등을 이용하여 최대발성지속시간을 알아보고, Aerophone-II(Kay Elemetrics Corps., USA) 등의 장비를 이용하여 성문하압(subglottal air pressure), 평균호기류율(mean airflow rate) 등을 관찰한다. 진단에 가장 유용한 검사로 알려진 것은 후두스트로보스코피인데, 특수내시경을 이용하여 성대의 병변이나 운동성, 진동의 패턴, 점막 파동 등을 영상으로 관찰한다. 이러한 평가들은 음성 전문 평가자의 기준에 따른 측정 방법으로서 음성장애의 종류와 정도를 판별하는 중요한 기준이 되어 왔다.

위의 평가들은 후두 질환에 대한 감별 및 음성 장애의 중증도를 파악하는 지표이지만, 평가자의 기준에서 시행한 것이므로 환자 본인이 주관적으로 느끼는 음성 문제의 정도를 반영하지는 못한다. 환자의 입장에서 자신의 음성을 평가할 때는 개개인의 상황, 예컨대 직업, 사회활동 등에 따라 음성 문제의 정도를 인식하는 수준이 상당히 다를 수 있다.⁷ 평가자 기준에서 여러 환자들이 같은 정도의 음성 문제를 보이는 것으로 판단되더라도 일상생활에서 음성 사용이 크게 중요하지 않은 특정 환자의 경우에는 특별히 불편감을 호소하지 않을 수 있다. 한편, 환자의 음성이 생계의 수단, 취미 활

동, 혹은 전문 직종에서 중요하게 사용되는 경우에는 자신의 음성 문제를 심각하게 여기고 불편을 호소한다. 또한 비슷한 정도의 음성 환자에게 같은 과정의 치료 중재를 제공하고도 환자에 따라서는 결과에 대한 기대치나 만족도가 매우 다를 수 있다. 따라서 음성장애는 환자 스스로가 음성의 문제 유무를 자가 보고하는 것이 평가에는 일차적 기준이 되고 치료에는 최종적 기준이 될 수 있다.

최근에는 환자가 일상생활에서 느끼는 음성 문제의 정도를 파악하여 이를 치료 계획에 반영하고, 치료의 호전도 및 만족도를 측정하는 도구로서 음성에 관련된 심리측정적 평가(voice-related psychometric evaluation)도구가 고안되어 임상뿐만 아니라 조사 연구에서도 그 사용이 현저히 증가하는 추세이다.⁷⁻¹² 또한 심리측정적 평가도구는 설문 조사 형태의 평가도구로서, 대상자가 병원이나 기관을 직접 방문하지 않아도 장애의 정도를 확인할 수 있는 장점이 있으므로, 대중적 조사 연구에도 용이하다.

그러나 우리나라에서는 최근에서야 음성장애에 관련된 심리측정적 평가도구를 사용하기 시작하였으므로 구축된 자료가 거의 없는 실정이다. 또한, 외국에서 이미 사용되어온 설문지를 번역하여 사용할 때도 번역본에 대한 신뢰도 및 타당도를 거친 공통 문항이 마련되지 않았다.

2. VHI(Voice Handicap Index) 및 기타 평가도구

음성장애에 관련된 심리측정적 평가도구 중에서 가장 대표적인 Voice Handicap Index(이하 VHI)는 1997년에 Jacobson 등(1997)에 의해 고안된 척도이다. 후두적출 환자를 포함하여 음성장애의 종류에 관계없이 성인 음성장애 환자를 대상으로 광범위하게 사용될 수 있다고 하였다.⁸ Jacobson 등은 면담을 통해 환자가 호소하는 음성 문제를 85개로 추렸고, 65명의 음성장애 환자에게 5점 척도로 평가하게 한 후, 각 문항에 대한 내적 일관성 신뢰도를 구하고, 최종적으로 신체적(P: physical), 기능적(F: functional), 그리고 정서적(E: emotional) 영역으로 나누어 각 하위 항목에 10개씩 총 30개의 문항을 구성하였다. 신체적 항목은 후두의 불편감과 발성의 문제

를 환자 스스로가 어떻게 인식하고 있는 지를 알아보는 것이고, 기능적 항목은 일상적 활동에서 음성 장애의 정도를 알아보는 것이고, 정서적 항목은 음성장애로 인해 겪게 되는 반응을 알아보는 것이다. 하위 영역을 위의 세 가지 영역으로 나눈 기준은 World Health Organization(WHO)¹³에서 헨디캡(handicap)을 손상(impairment)이나 무능력(disability)으로 인해 사회적, 경제적, 환경적 측면으로 불이익을 보이는 것으로 정의한 것에 기초하여 연구자들에 의해 임의적으로 내용을 구분한 것이었다.

VHI는 음성장애 환자를 대상으로 삶의 질(quality of life)을 알아보기 위한 설문으로 사용되거나, 치료 효과를 알아보기 위해 치료 전후 비교 연구 등에 사용되어왔다.¹⁴⁻²¹ 언어병리분야에서 사용되고 있는 평가들을 조사한 Agency for Health care Research and Quality에 의하면 음성장애와 관련된 도구 중에서 VHI가 유일하게 신뢰도와 타당도 기준에 부합된다고 발표한 바가 있었기에,²² 현재까지도 이를 사용한 연구는 계속 증가하고 있다.

그러나 VHI를 표준화된 평가도구로 사용하는 것은 아직 문제가 있음을 지적하기도 하였다. VHI를 제작하는 과정에서 실험 대상자의 수가 적어서 타당도가 떨어질 수 있고, 세 가지의 하위 항목에 대한 요인분석(factor analysis)을 실행하지 않고 임의로 설정한 것이어서 그 구분이 명확하지 않다는 점,^{7,23} 30개의 문항이 임상에서 실제로 사용하기에는 문항수가 많아 시간적인 제약을 받는 등을 지적받기도 하였다.^{7,12} 또한 전문적인 음성 사용자를 포함하여 다양한 직업에 관계없이 사용할 수 있다는 보고와는 달리, Rosen 등(2004)의 연구에서는 음성에 문제가 있는 환자를 대상으로 노래 전문인과 일반인을 비교하여 VHI를 실시한 결과, 노래 전문인에 비해 일반인에서 총점이 유의하게 증가되어 직업에 따른 차이가 있음을 지적하였다.²⁴ 이러한 현상은 VHI 문항들이 주로 말 수준에 대한 문항들로서, 노래 수준에서 나타나는 음성 문제를 반영하는 문항이 없었던 것을 원인의 하나로 설명하였다. 위와 같이 VHI에 대해서는 일부 문제점이 지적되고 있지만, 현재까지 가장 많이 사용되어온 도구로서 연구 결과 보고가 상대적으로 많아 기존 연구들과 비교가 용이하다는 장점이 있다.

VHI 외에도 이와 유사한 음성장애관련 심리측정적 평가도구들이 많이 소개되어 왔는데, 평가의 성격 및 적용 대상에 따라 다양한 문항과 측정 방법을 제시해 왔다. 미국의 사회 보장 관리청(Social Security Administration, SSA)에서는 음성장애의 현

황을 파악하는 심리평가도구를 선정하기 위해 9개의 평가 도구 즉, VHI,^{8,15,20} VAPP(Voice Activity and Participation Profile),^{25,26} V-RQOL(Voice-Related Quality of Life) measure,^{9,27,28} VCQOLQ(Voice Clinic Quality of Life Questionnaire),²⁹ VDI(Voice Disability Index), VOS(Voice Outcome Survey), Pediatric VOS(Pediatric Voice Outcome Survey), HNQOL(University of Michigan Head and Neck Quality of Life) instrument,³⁰ 그리고 UW-QOL-R(University of Washington Quality of Life Questionnaire Revised)³¹에 대한 신뢰도 및 타당도와 질병 예견력을 문헌 조사를 통해 알아보았다.³² 문헌 조사에 대상이 된 논문들 중 VHI가 인용된 논문이 19편으로 가장 많았고, 나머지의 평가 도구들은 논문이 없거나 1편에서 최대 4편 정도였다. 조사 결과, 아직은 표준화되거나 공통적으로 사용될 수 있는 도구가 없으며, VHI는 타당도가 결여되어 있어 개인을 대상으로 한 연구로 사용을 제한하고, 그룹 간 비교 연구에는 V-RQOL가 더 적합하다고 보고하였다.³² 1999년에 제작된 V-RQOL⁹은 음성장애와 연관된 10개의 문항으로 구성되어 있고, 두 가지의 영역, 즉, 사회-정서적인 면과 신체-기능적인 면으로 구분한다. 하지만 V-RQOL은 VHI에 비해서 문헌에서의 이용도가 상대적으로 낮은 편이어서, 현재로서는 비교 연구가 적절하지 않은 실정이다.

이 밖에도, VoiSS(Voice Symptom Scale)^{10,23}는 2003년도에 Deary 등에 의해 발표된 것으로, 총 43개의 문항 중 13개는 VHI와 같은 내용을 사용하였고, 음성 문제 외에 동반되는 ‘목 안의 통증’, ‘기침 및 목 가다듬기’ 등 전반적인 후두 증상과 일상 발화 시 나타나는 음성 문제 외로 ‘노래 시 나타나는 발성 문제’에 대한 내용도 포함되었다. 또한, VPQ(Vocal Performance Questionnaire)^{11,21}는 영국에서 주로 사용하는 것으로서, 비기질적인 음성장애를 대상으로 음성치료 효과를 보기 위해 고안된 12개의 문항으로서 음성 문제의 생리적인 측면과 사회 및 정서적인 측면이 설문 내용에 구성되어 있다.

최근에는 임상 현장을 고려하여 기존에 소개된 음성장애관련 심리측정적 평가도구들에 비해 문항 수를 줄여 설문 시간을 단축하면서 효용성 있는 문항 구성을 제시하기도 하였다. GFI(Glottal Function Index)¹²는 생리적 기능에만 국한하여 총 4개 문항으로서 성대 수준에서의 이상 증상을 측정한다. IPVI(Iowa Patient’s Voice Index)²⁾

는 미국 아이오와 대학 Verdolini 연구소에서 사용하는 척도로서 소개된 바 있고, 음성에 관련된 총 3개의 문항으로 구성되어, 환자 스스로가 자신의 음성(음질)에 대한 인식 정도, 음성 사용 시 노력 정도, 자신의 삶(생활)에서 음성의 중요도 정도를 측정한다.

3. 공인타당도 검증 문제

현재까지는 표준화된 음성장애관련 심리측정적 평가도구는 없으므로, 그간에 제작된 도구들의 공인타당도(concurrent validity) 평가는 총체적인 음성 문제에 관한 질문에 대한 반응 정도와 상관성을 보는 것으로 대신하고 있다. VHI를 제시한 연구자들은 피험자에게 본인의 음성 문제가 어느 정도인지를 평가하게 하는 한 가지의 질문을 제공하고, 이를 4점 척도(0: 정상, 1: 경도의 문제, 2: 중도의 문제, 3: 심도의 문제)로 정도를 측정하여 타당성 평가에 사용한 바 있다.⁸

V-RQOL를 제시한 연구자들은 일반적인 삶의 질을 평가하는 Medical Outcomes Study Short Form 36-Item Health Survey(SF-36)³³⁻³⁷을 비교하는 것으로 대신하였다.⁹ SF-36은 음성장애에 관한 특화된 설문지가 아니고, 일반적인 건강을 통해 삶의 질을 측정하는 공인된 검사지이다. 저자들은 음성장애로 인한 삶의 질과 일반적인 건강에 대한 삶의 질을 비교하여 일차적으로 정상군과 음성장애군 간의 차이를 비교하고, 각 검사지의 하위 항목을 비교하여 타당성을 보고자 하였다. Cohen 등³⁸은 VHI와 SF-36에 대한 문헌 조사를 하였고, 후두암을 제외한 음성장애를 대상으로 정상군과 장애군 간 비교 및 질환별 장애군에 대한 비교를 통해 설문의 타당성을 조사한 바 있다.

최근의 연구에서는 기존 연구에서 사용 빈도가 높았던 VHI나 기타의 음성장애관련 심리측정적 평가도구와 비교하여 타당성을 측정하기도 한다.^{11,23} 그러나 이는 단순히

2) Karnell MP, Melton SD, Childes JM, Coleman TC, Dailey SA, Hoffman HT. Reliability clinician based(GRBAS and CAPE-V) and patient-based(V-RQOL and IPVI) documentation of voice disorders. J Voice Accepted for publication May 2, 2006.

두 검사도구 간의 상관성을 알아본 것이고, 표준화된 검사를 기준하여 공인타당도를 측정된 결과로 해석하기는 어렵다. 이를 만족하기 위해서는 VHI를 포함하여 음성장애 관련 심리측정적 평가도구에 대한 표준화 작업이 선행되어야 할 것이다.

4. 질환군별 VHI 결과 비교

VHI는 환자에 의해 음성장애의 유무와 그 정도를 주관적으로 평가하는 것으로, 총점 0점부터 120점까지 VHI의 점수를 통해 일차적으로 정상과 음성장애를 구분해볼 수 있다. Cohen 등(2006)이 VHI에 대한 27개의 문헌을 조사한 결과에서는 정상군의 VHI 평균 총점이 5.3이었고, 장애군의 질환별 VHI 평균 총점은 대표적으로 경련성 발성장애(spasmodic dysphonia)가 73.0, 편측성대마비(unilateral vocal fold palsy)가 68.1, 성대폴립(vocal polyp)이 52.1, 라인케씨부종(Reinke's edema)이 42.9, 성대결절(vocal nodules)이 33.4이었다.³⁸ Rosen의 연구(2004)에서는 정상군의 VHI 총점 평균(±표준편차)이 8.75(±14.97)이었고, 중위수 값은 4.0이었다. 장애군의 질환별 총점 평균은 대표적으로 경련성발성장애가 74.30(±23.09), 성대마비가 69.47(±27.21), 라인케씨부종이 47.26(±30.11), 성대폴립/성대결절/성대낭종(vocal cyst)이 36.96(±19.80)이었다.⁷ 또한 Rosen 등은 일차 연구에 참여한 소수의 환자군과 이차 연구에 참여한 다수의 환자군에 대한 VHI의 총점을 비교한 결과, 성대마비의 경우 75점(14명)과 69점(104명)을, 성대폴립/성대결절/성대낭종의 경우 41점(13명)과 37점(166명)을 보이는 등, 비교적 자료의 일관성을 보여 차후 통합연구에 공통 자료로서 사용 가능성을 시사하기도 하였다. 이는 SSA³²에서 VHI가 제작 시 타당성 검증이 부족했던 이유로 그룹 간 비교 연구에는 사용 제한이 있음을 보고한 것과 달리, 향후 자료 구축을 통해 사용 가능성을 확인해 볼 필요가 있음을 보였다.

음성장애의 하위 질환 구분은 연구자마다 설정 기준이 다르기 때문에 개별 질환에 대한 결과 값을 대응하여 비교하기는 어려울 수 있다. 예를 들어, 기능적 장애의 경우, 성대의 병변을 동반하지 않은 근긴장성발성장애 등을 국한하여 지칭하는 경우도 있

지만, 연구자에 따라서는 발병 원인이 기능적인 이유인 성대폴립 또는 성대결절 등도 이에 포함하는 경우가 있기 때문이다. 그러나 질환 구분의 모호함에도 불구하고, 위의 선행연구들을 통해 적어도 신경학적 장애에 해당되는 경련성발성장애 및 성대마비가 기타 구조적 또는 기능적 음성장애에 해당하는 성대폴립이나 성대결절에 비해 VHI 총점이 뚜렷이 높음을 알 수 있다. VHI가 환자 개인의 주관적 평가임에도 불구하고, 음성 문제로 인한 삶의 질은 음성장애 질환군에 따라 정도에 차이가 있다.

5. VHI와 청지각적/음향학적 검사 결과 간의 비교

설문을 통한 평가는 환자의 주관성을 반영한 결과를 얻는 것이므로, 평가자의 기준에 의한 검사 결과와 어떤 상관성을 보이는지 알아볼 필요가 있다. 일반적으로 훈련된 평가자에 의한 주관적 평가와 객관적 평가는 상관성이 있다. 예컨대, 청지각적 평가인 GRBAS와 음향학적인 평가의 파라미터들 간에는 비교적 상관성이 있는 것으로 알려져 있다.³⁹ 또한, GRBAS와 후두스트로보스코피의 후두 소견 간에도 상관성이 있음을 보고한 바 있다.⁴⁰ 결국 훈련된 평가자에 의한 주관적인 평가는 평가의 척도로서 비교적 일관성을 보인다고 설명할 수 있다.

기존 연구에서 VHI와 음향학적인 수치 간에는 상관성이 거의 없는 것으로 보고된 바가 있다. Wheeler등의 연구³⁾에서는 음향학적 평가 중 jitter, shimmer, SNR(signal to noise ratio) 등을 VHI의 총점과 개별 문항 점수에 비교했을 때 상관성이 적다고 보고하였다. 또한, 치료 전에 VHI를 GRBAS 및 음향학적 평가에 비교했을 때 높은 상관성을 보였었는데, 음성 치료 후에는 상관성이 낮아졌으므로, 치료 효과에 대한 환자의 주관적 견해가 객관적 수치와 다를 수 있음을 시사하였다.⁴¹ 즉, 치료 후 청지각적 평가와 음향학적 평가의 수치는 호전되어도 환자의 주관적인 호전도에는 큰 변화가 없을 수 있고, 그와 반대로, 청지각적 평가나 음향학적 평가의 수치에는 뚜렷한 변화가 없더라도 환자의 주관적인 호전도는 증가되는 경우가 있을 수 있다.

3) Wheeler KM, Collins SP, Sapienza CM. The relationship between VHI scores and specific acoustic measures of mildly disordered voice production. J Voice 2005; in press.

6. VHI-10의 유용성

시간적 제약을 극복하기 위해, Rosen 등(2004)은 신뢰도 및 타당도를 유지하면서 VHI 30개 문항(이하 VHI-30)을 하위 영역의 구분 없이 10개의 문항(이하 VHI-10)으로 줄였다.⁷ 선정 기준은 환자를 대상으로 치료 전후에 가장 편차를 많이 보인 문항, 음성장애인과 정상 대조군 간에 차이를 많이 보인 문항, 그리고 전문가의 종합적 의견에 의한 추천 등이었다. 전문가가 추천한 항목은 수입(income)에 대한 항목으로, 위의 두 조건에는 해당되지 않았지만 헨디캡을 나타내는 항목으로서 필요하다는 내용타당도에 의한 추천이었다. VHI-10과 VHI-30을 비교한 결과, 이 두 검사 간의 상관성이 매우 높았으며, 모든 후두 질환에서 VHI-10 평균 총점 값 대 VHI-30 평균 총점 값의 비율이 0.333보다 높아, VHI-10의 민감도가 높음을 증명하기도 하였다.

그러나, 현재까지는 VHI-10을 이용한 연구가 많지 않아서, 신뢰도나 타당도에 대한 적용 결과를 얻지 못하고 있다. VHI-10에 대한 연구는 중국어⁴²와 히브리어⁴³에서 번안본 VHI-10에 대한 타당성을 알아본 것이 있다.

VHI-10은 VHI-30에 비해 사용 간편성이 가장 큰 장점이 되므로, 앞으로 임상에서 사용 빈도가 늘어날 것으로 보인다. 그러나 VHI-10의 본격적인 사용에 앞서 풀어야 할 과제가 있는데, 첫째, Rosen 등이 주장한 바와 같이 과연 VHI-30에 대한 세 가지 하위항목 구분이 평가에 타당성이 없는 것인지와, 둘째, VHI-30 중 채택된 VHI-10 문항을 다른 언어 및 문화권에서도 마찬가지로 타당성이 높은 문항으로 선택할 수 있는지를 알아보는 것이다.

7. VHI 번안본의 검증 필요성

VHI는 영어뿐만 아니라 프랑스어,⁴⁴ 네덜란드어,⁴⁵ 포르투갈어,⁴⁶ 히브리어,⁴³ 대만어,⁴⁷ 중국어⁴² 등 다수의 언어로 번안되어 사용되고 있고, 각 언어권에서 신뢰도 및 타당도에 대한 연구가 진행 중이다. 일반적으로 외국어로 된 설문지 문항을 번안하여

사용할 때는 언어적 차이와 문화적 차이를 잘 반영했는지를 고려해야 한다.⁴⁸ 국내에서는 기관마다 자체 번안된 설문지를 사용하고 있는데,^{49,50} 이들 간에 번역 자체의 질에 대한 확인이 필요하며, 문화적 타당성도 검증을 하여야 한다.

번역된 문항은 언어의 차이로 인해 원본의 뜻이 완전히 동일하게 전달되지 않을 수 있다.⁴⁸ 따라서 원본의 본래 의미를 왜곡하지 않으면서 해당 언어에 맞는 적절한 번안이 필요하다. 또한 설문에서 사용되는 문항들은 일반인이 쉽게 이해할 수 있는 문장으로 구성되어야 한다. 번역한 문항이 어색하거나 흔히 사용하는 용어가 아닌 경우에는 피검자로부터 적절한 반응을 기대할 수 없고, 피검자가 설문 시행 중에 제공자에게 추가 설명을 요구할 수 있어 검사 중 간섭 효과가 나타날 수도 있기 때문이다.

VHI를 타문화의 다른 언어로 번안하여 사용하는 연구자들은 모국어 수준의 이중 언어 사용자에게 이를 확인하는 작업을 거치거나, 영어를 해당 언어로 번역한 후 번역본을 다시 영어로 역번역을 하여 처음 문장과 비교해보는 이중번역을 하기도 한다. 또한 검사 타당화 검증을 거쳐 번안된 설문지의 완성도를 높인다.⁵¹

논문을 통해 소개된 VHI 한국어 번역본은 최두영 등(2002)⁴⁹과 박성신(2004)⁵⁰의 연구에서 이미 제시된 바 있다. 최두영의 연구에서는 번역본에 대해 검사-재검사에 대한 신뢰도만을 제시하였고, 타당도는 검증하지 않았다.

8. 문화적 차이를 반영하는 문항의 필요성

문화적 차이에 대한 연구는 여러 분야에서 다루어져 왔고, 설문지의 사용에서도 이에 대한 고려가 요구된다. 병원 기관에서 사용하는 설문지나 증상 측정 검사지들에 대해 대상자의 문화 배경에 따른 차이가 있는지를 조사한 연구 결과에서는, 대부분의 항목은 공통 적용이 가능하나 일부 문항에서는 차이가 있었음을 지적한 바 있다.⁵² 따라서 환자가 일상생활에서 활동을 하면서, 음성으로 인해 느끼는 문제를 측정하고자 한다면, 환자가 속해 있는 사회 문화 및 정서를 반영하는 문항으로 재구성을 해보는 것도 필요하다. 음성장애를 주소로 내원한 환자를 면담할 때, 말 수준뿐만 아니라 노래 수준에서의 음성 사용에 대한 불편감을 호소하는 경우를 흔히 접하게 된다. 직업적으

로 노래를 부르는 사람이 아닌 일반인의 경우에도 우리나라의 사회 문화적 활동 상 노래에 대한 요구가 많은 편이다. 환자에 따라서는, 말을 할 때는 목소리 문제를 느끼지 못하지만 노래를 부를 때는 고음 발성에서 제한이 있다거나, 큰소리가 나오지 않거나, 목에 통증을 느낀다는 등의 증상을 보고하기도 하며, 노래방에서 사람들과 어울려 노래를 부르지 못하는 것에 대한 스트레스가 있다고 보고 하기도 한다. 이러한 문제 호소가 음성장애를 반영하는 평가 문항에 포함된다면 그 타당도를 보아야 할 것이고, 또 포함이 되었다면 전체 문항에서 어느 정도의 비중을 차지하는지에 대한 분석이 선행되어야 할 것이다.

9. 연구의 목적

음성장애관련 심리측정적 평가도구로 가장 보편적으로 많이 사용되어 온 VHI-30 및 VHI-10을 한국어로 번안한 한국어판 VHI의 신뢰도와 타당도를 검증하고자한다. 또한, 질환군에 따른 VHI-30의 결과 비교 및 VHI-30과 객관적 검사 결과 간의 상관성을 확인하여, 음성장애 평가 도구로서 VHI의 유용성을 알아본다.

실문에 가장 효과적인 문항 구성을 제시해 보고자, 기존의 VHI-10과 본 연구에서 VHI-30 중 새로 선택한 10개 문항을 비교하여 민감도를 비교하고자 한다. 또한 VHI-30 문항 외에 한국의 문화적 특성을 반영한 ‘노래 부르기’ 문항을 첨부하여, 그 타당성을 확인한다.

이에, 본 연구의 세부 목적은 다음과 같다.

첫째, 한국어로 번안된 VHI-30의 신뢰도 및 타당도를 검증한다.

신뢰도 검증으로,

(1) 환자군 일부에 대한 검사-재검사(test-retest)로 피험자 내 일치도 (intraclass correlation coefficient)를 알아본다.

(2) Cronbach's α 값으로 문항내적일관성(internal consistency)을 알아본다.

(3) 각 문항과 총점 간의 상관성으로 문항변별도를 알아본다.

타당도 검증으로,

(1) 전문가에 의한 내용타당도(content validity) 검증단계를 거쳐 한국어판 VHI-30을 제시한다.

(2) 환자군과 정상대조군을 비교하여, 음성장애군과 정상군에 대해 구별 가능한 평가도구임을 확인한다.

(3) VHI-30의 총점과 IPVI의 1번 문항을 비교하여 공인타당도를 대체하여 알아본다. IPVI의 1번 문항은 총괄적으로 음성 문제 정도를 알아보는 것으로, '내 목소리에 문제가 없다'(0점)부터 '내 목소리에 문제가 심각하다'(6점)에 대해 7점 척도로 정도를 표기한다.

(4) VHI-30과 SF-36의 총점 및 하위항목을 비교하여, 일반적인 건강에 관련된 삶의 질을 측정하는 공인 검사지를 기준한 공인타당도를 알아본다.

(5) 요인분석(factor analysis)을 하여 VHI-30의 기존 하위 항목인 기능적 항목, 신체적 항목, 정서적 항목에 대한 구인타당도(construct validity)를 알아본다.

둘째, 환자군의 질환군(신경학적 장애, 기능적 장애, 구조적 장애, 염증성 장애)에 따라 VHI-30의 총점을 비교하여, 만약 차이가 있다면, 질환군별 비교연구에 VHI-30의 사용 가능성을 제시할 수 있다. 본 연구에서는 Boone(2005)의 질환 분류¹에 기초하여 음성장애 질환군을 구별하였다. 따라서 기능적 장애에는 근긴장성발성장애뿐만 아니라 과다기능이 원인이 된 성대폴립이나 성대결절 등이 해당된다.

셋째, 환자군을 대상으로 VHI-30과 음향학적 평가 결과 간의 상관성을 보고, 환자의 주관적 평가와 평가자의 객관적 평가 간에 차이점이 있는지 알아본다. 상관성이 있다면, VHI-30이 객관적인 음성 평가에 상응하는 동일 차원의 평가로 해석할 수 있고, 반대로 상관성이 없다면, 객관적인 평가와는 차원이 다른 독립적인 평가로 해석할 수 있다. 본 연구에서는 VHI와 GRBAS 간의 비교는 제외하였는데, 기관별 GRBAS의 기준에 대한 평가자 간 차이를 고려하였기 때문이다.

넷째, VHI-30과 VHI-10을 비교하여 VHI-10의 임상적 유용성을 재확인한다. VHI-10 평균 총점 값 대 VHI-30 평균 총점 값 비율이 0.333보다 높은 경우, 단축된 문항으로서 평가에 효율성을 보인 것으로 해석한다.

다섯째, VHI-30에서 10개의 문항을 새로 선별하여(가칭 new VHI-10), 기존의 VHI-10과 비교한다. 문항 선별 기준은 VHI-30 중 정상군과 환자군의 평균차를 많이 보인 문항, 삭제 시 전체 Cronbach's α 가 증가하는 문항, 총점에 대한 상관계수가 높은 문항이 해당된다.

여섯째, VHI-30 문항에 문화적 특성을 반영한 문항을 추가하여, 이 중 10개의 문항을 새로 추려서(가칭 singing VHI-10), 기존의 VHI-10과 비교한다. 본 연구에서는 문화적 특성을 반영한 내용을 '노래 부르기에 대한 문항으로 국한하여, VHI의 하위항목에 따라 신체적인 측면에서, '노래 부르기가 힘들다', 기능적인 측면에서, '사람들과 어울릴 때 노래를 잘 부르지 못해 불편할 때가 있다', 그리고 정서적인 측면에서, '누가 나에게 노래를 시킬까봐 걱정을 할 때가 있다'로 구성하였다. 문항 선별 기준은 VHI 30문항과 '노래 부르기' 세 문항을 포함한 총 33개의 문항 중 정상군과 환자군의 평균차를 많이 보인 문항, 삭제 시 전체 Cronbach's α 가 증가하는 문항, 그리고 총점에 대한 상관계수가 높은 문항이 해당된다.

II. 대상 및 방법

1. 연구 대상

실험군인 음성장애 환자군과 대조군인 정상군은 모두 20세 이상 남녀 성인으로, 제공된 설문지에 대한 한글 독해 및 수행이 가능한 자로 제한하였다. 대상자는 정신장애 또는 신체장애를 동반하지 않았다.

가. 환자실험군

음성장애 환자군은 음성 문제를 주소로 이비인후과에 내원하여 전문의의 진단을 받고 음성언어치료사가 음성평가를 시행한 경우로서, 후두암을 제외한 음성 장애 환자들 중 설문 전에 수술적인 치료 또는 음성치료 등의 중재가 제공되지 않는 상태에서 설문을 시행한 156명이 대상이 되었다. 이들은 2007년 3월부터 4월 사이에 삼성서울병원, 영동세브란스병원, 서울아산병원, 서울대학병원, 아주대학병원, 인하대학병원 등의 이비인후과 음성언어치료실에 내원한 환자를 대상으로 수집되었다.

환자군의 연령은 20세부터 80세까지로 분포되었고, 평균 45.0세(표준편차: 13.0)이었다. 성별 분포는 남자 68명(43.6%) 여자 88명(56.4%)이었다(표 1). 환자의 음성 질환은 다양하여, 성대마비(vocal fold palsy) 16명, 경련성 발성장애(spasmodic dysphonia) 5명, 성대폴립(vocal polyp) 45명, 성대결절(vocal nodules) 22명, 라인케씨부종(Reinke's edema) 7명, 근긴장성발성장애(muscle tension dysphonia) 6명, 성대구증(sulcus vocalis) 5명, 성대근위축(vocal atrophy) 2명, 후두유두종(laryngeal papilloma) 4명, 성대낭종(vocal cyst) 6명, 피열연골골절(arytenoid cartilage fracture) 1명, 후두격막(laryngeal web) 1명, 만성후두염(chronic laryngitis) 17명, 후두인두역류염(laryngopharyngeal reflux) 14명, 위산역류에 동반된 접촉성궤양(contact ulcer) 및 육아종(contact

granuloma) 4명이었다(표 2). 환자가 호소하는 음성변화의 유지 기간은 1개월 미만인 9명, 1개월 이상부터 6개월 미만이 61명, 6개월 이상부터 1년 미만이 18명, 1년 이상부터 3년 미만이 28명, 3년 이상 10년 미만이 27명, 10년 이상이 13명이었다. 직업은 다양하여, 교사, 목사, 가수, 성악가 등 전문적 음성 사용자가 23명이었고, 직업적 음성 사용이 많은 판매직, 상업 종사자 등은 17명이었고, 비교적 직업적 음성 사용이 보통 정도인 일반 사무직, 자영업 등은 32명, 주부 13명, 학생 4명, 무직 5명으로 구성되었다.

나. 정상대조군

정상군은 평소 음성 문제를 특별히 느끼지 않았고, 최근 3개월 내에 음성 문제로 인해 이비인후과에 내원한 적이 없는 자로, 제공된 IPVI 1번 문항(부록 1)에서 해당 값을 '0' 또는 '1'로 기록을 한 66명을 대상으로 하였다. 해당 값이 '2' 이상인 경우는 정상군에서 제외하였다. IPVI 1번 문항에서 '2' 이상으로 기록한 경우는 교사, 목사 등 전문적 또는 직업적 음성 사용이 많은 직업인이 대부분이었다.

정상군의 연령은 20세부터 75세까지 분포되었고, 평균 45.4세(표준편차: 15.5)이었고, 성별은 남자 33명(50%), 여자 33명(50%)이었다(표 1). 직업은 교사, 목사, 가수, 성악가 등 전문적 음성 사용자가 7명이었고, 직업적 음성 사용이 많은 판매직, 상업 종사자 등은 5명이었고, 비교적 직업적 음성 사용이 보통 정도인 일반 사무직, 자영업 등은 32명, 주부 13명, 학생 4명, 무직 5명으로 구성되었다.

표 1. 환자군과 정상군의 연령 및 성별 분포

	환자군	정상군
연령(세)		
범위	20~80	20~75
평균(±표준편차)	45.0(±13.0)	45.4(±15.5)
성별(명)		
남(%)	68(43.6)	33(50)
여(%)	88(56.4)	33(50)
총계	156	66

표 2. 환자군의 음성장애별 분포

구분	질환명	명수
신경학적 장애	성대마비	16
	경련성발성장애	5
기능적 장애	성대폴립	45
	성대결절	22
	라인케씨부종	7
	근긴장성발성장애	6
구조적 장애	성대구증	5
	성대근위축	2
	후두유두종	4
	성대낭종	6
	피열연골골절	1
	후두격막	1
염증성 장애	만성 후두염	17
	후두인두역류염	14
	접촉성궤양/육아종	4
총계		156

2. 연구 방법

가. 한국어판 VHI 제작

기존에 이미 한국어로 번안되어 논문에 인용된 바 있는 2개의 번안본^{30,31}과 한국어와 영어의 모국어 수준 이중언어 사용자 2명에게 감수를 받은 새 번안본 1개를 준비하였다. 그리고 적어도 5년 이상의 음성장애 평가 및 치료 경험이 있는 5명의 음성언어치료사 및 1명의 연구원, 총 6명의 전문가에게 내용타당도 평가를 의뢰하였다. 각 기관별 참여자 수는 삼성서울병원, 신촌세브란스병원, 강북삼성병원, 아주대학병원 각 1인씩, 그리고 영동세브란스병원 2인이었다.

내용타당도를 평가함에 있어서, 각 평가자는 번역된 문장이 영어 원문장 뜻에 충실하며, 임상에서 환자 면담시 흔히 사용하게 되는 문구로서 내용전달이 비교적 명료하고 자연스러운 지에 대해 5점 척도(1: 상당히 부적절, 2: 약간 부적절, 3: 중간(보통) 정도, 4: 비교적 적절, 5: 상당히 적절)로 점수를 매겼다. 또한, 표현에 수정을 요하는 경우에는 피험자가 문장을 새로 작성해보도록 하였다. 30문항의 각 문항에 대해 세 가지 번역본 중 한 번역본에서 3명 이상의 평가자가 5점을 부여한 18개의 문항들을 일차로 채택하였고, 한 문항에서 5점이 3명 미만을 보인 기타 문항들은 평균값이 가장 높은 문장을 채택하였다. 또한, 피험자가 문장을 번역한 것을 참조하여 일부 수정한 경우(예, ‘얼굴을 맞대고’를 ‘얼굴을 마주보고’로 수정)에는 2명의 음성언어치료사와 1명의 이중언어 사용자의 재확인 절차를 거쳤다.

완성된 VHI 번안본은 본 실험의 실행 이전에 20명의 음성장애 환자에게 무작위로 제공하여 시범적으로 실시해보았다. 설문에 응하는 과정에서 문항에 대한 특별한 질문 없이 혼자서 수행이 가능함을 확인하였다.

나. 설문지 구성

내용타당도 검증 절차를 거친 VHI-30은 Jacobson이 제시한 문항 순서에 따라 제공되었다(부록 2). 일부 선행 연구에서는 검사자의 사용 간편성을 위해 기능적인 측면, 신체적인 측면, 정서적인 측면을 각각 분류하여 제공하기도 하였으나, 이는 비슷한 질문의 내용이 10개 문항씩 연속되는 구성에서 야기되는 오류를 배제할 수 없다. 또한 원본과 같은 조건에서 설문지의 신뢰도 및 타당도를 비교해 보기 위해서 이를 준수하는 것이 요구되었다. 또한, 기존 VHI 30문항에 연속하여 '노래 부르기'에 해당하는 3개의 문항을 첨부하였다. 척도는 VHI와 마찬가지로 0점부터 4점까지의 5점 척도로 정도를 표기하게 하였다(부록 3). VHI와의 타당성을 보기 위해 한국어 번안본 IPVI 1번 문항(부록 1), SF-36 문항도 동시에 제공되었다(부록 4). SF-36은 기존 번안본³⁷을 참고하여 재 번역을 한 후 이중언어 사용자의 감수를 받았다.

다. 설문지 시행 및 수집

(1) VHI, IPVI 및 SF-36 시행

대상자에게 VHI-30 및 '노래 부르기' 세 문항과 IPVI 1번 문항, SF-36을 무작위로 작성하게 하였다. 일반적으로 피험자가 설문을 완성하는 하는 시간은 대략 10분 내외였다. 설문을 거부 또는 도중에 중단한 경우는 대상에서 제외하였다. 또한 표기가 누락된 문항은 검사자가 그 자리에서 확인을 하여 표기를 요구하거나 피험자에게 전화 통화로 해당 값을 얻었지만, 이미 수거한 후 3일 이상이 지나 누락 문항이 발견된 설문은 시간에 따른 조건 변화가 변수가 될 수 있으므로 대상에서 제외하였다. 최종적으로 환자군 156명과 정상군 66명의 설문지가 사용되었다.

(2) 검사-재검사

환자군 156명 중 약 10%에 해당되는 18명에게 VHI-30에 대한 검사-재검사를 실시하였다. 재검사는 1차 검사일로부터 1일 이상 1주 이내에 재 진료 방문 시 시행되었다. 재검사가 1일 이상 경과 후 시행된 것은 피검자 간 기억력 차이를 배제하고자

하는 것이었고, 1주 이내에 시행된 것은 기간의 차이에 따른 증상의 변화를 최소화하기 위한 것이었다.

(3) VHI-10 시행

VHI-30에서 추출한 VHI-10에 해당하는 10 문항과, VHI-10으로만 구성된 설문 간의 일치도를 알아보기 위해, 환자군의 약 10%에 해당하는 15명에게 무작위로 VHI-10 설문(부록 5)을 추가하여 제공하였다. 기존의 연구에서는 VHI-30과 VHI-10을 함께 제공하여 비교한 연구도 있었고,²⁸ VHI-30만을 제공한 후 전체 30문항과 VHI-10에 해당하는 10문항을 추출하여 비교하기도 하였다.²⁷ 본 연구에서는 일치도가 높을 경우, VHI-30에서 추출한 VHI-10로 결과를 제시하고자 한다.

라. 음향학적 검사

환자군을 대상으로 모음에 대한 음향분석은 MDVP로 실시하였고, 평소 대화와 같은 정도의 음도와 강도의 목소리로 모음 /아/를 지속 발성하게 하여 약 3초 내외의 안정된 구간을 취해 분석하였다.⁵⁴ 자동적으로 분석이 제공된 파라미터의 결과들 중, 주기 간 주파수 변동률을 보는 jitter와 주기 간 진폭 변동률을 보는 shimmer, 잡음 대 배음 비율을 보는 NHR 값을 구하였다.⁵³

3. 통계 분석

SPSS(Statistics Package for the Social Science, version 12.0, USA)로 통계 분석을 실시하였다. 정규분포를 이루지 않은 경우에는 비모수적 통계를 적용하였다.

Ⅲ. 연구 결과

1. VHI의 신뢰도 및 타당도

가. VHI 결과 분석

환자군 156명에 대한 각 문항의 평균과 표준편차는 표 3과 같다. VHI-30의 총점의 분포는 2점부터 115점까지로 비교적 정규분포를 이루었고, 평균은 49.61점(표준편차: 28.403)이었다(그림 1).

정상군 66명에 대한 각 문항의 평균과 표준편차는 표 3과 같다. VHI-30의 총점의 분포는 0점부터 34점까지였고, 평균은 5.18점(표준편차: 6.832)이었다. 0점 쪽으로 치우친 비정규분포를 이루었고, 중위수는 2.5점(사분위수 25th:1, 75th:8)이었다(그림 2).

표 3. 환자군과 정상군의 VHI(Voice Handicap Index)-30 문항별 평균과 표준편차

문항	환자군(156명)		정상군(66명)	
	평균	표준편차	평균	표준편차
F1	1.65	1.308	0.30	0.554
P2	1.40	1.308	0.12	0.541
F3	1.85	1.400	0.56	0.787
P4	2.33	1.108	0.42	0.658
F5	1.24	1.188	0.18	0.493
F6	1.79	1.362	0.18	0.461
E7	1.67	1.411	0.11	0.310
F8	1.10	1.254	0.08	0.267
E9	1.32	1.260	0.20	0.437
P10	2.37	1.171	0.14	0.426
F11	1.51	1.327	0.05	0.210
F12	1.17	1.217	0.21	0.481
P13	2.79	1.136	0.24	0.556
P14	2.21	1.309	0.21	0.512
E15	1.49	1.215	0.09	0.290
F16	1.63	1.442	0.11	0.310
P17	2.44	1.401	0.17	0.414
P18	1.90	1.333	0.26	0.535
F19	0.96	1.171	0.06	0.240
P20	1.81	1.376	0.08	0.267
P21	2.00	1.376	0.42	0.609
F22	1.00	1.305	0.11	0.310
E23	2.47	1.389	0.12	0.373
E24	1.79	1.501	0.12	0.412
E25	1.77	1.489	0.06	0.240
P26	1.77	1.377	0.18	0.461
E27	1.15	1.192	0.12	0.329
E28	0.85	1.092	0.11	0.310
E29	1.25	1.371	0.11	0.310
E30	0.92	1.161	0.08	0.267
총점	49.61	28.403	5.18	6.832

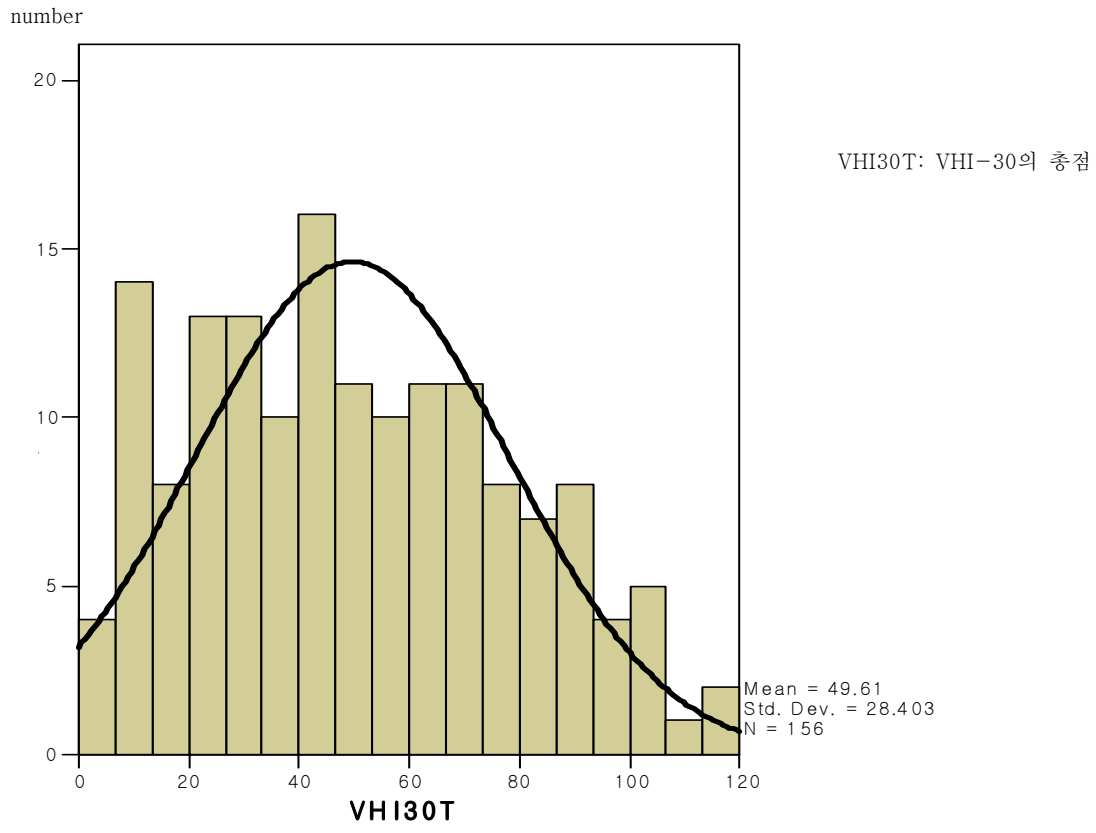


그림 1. 환자군의 VHI(Voice Handicap Index)-30의 총점 분포.

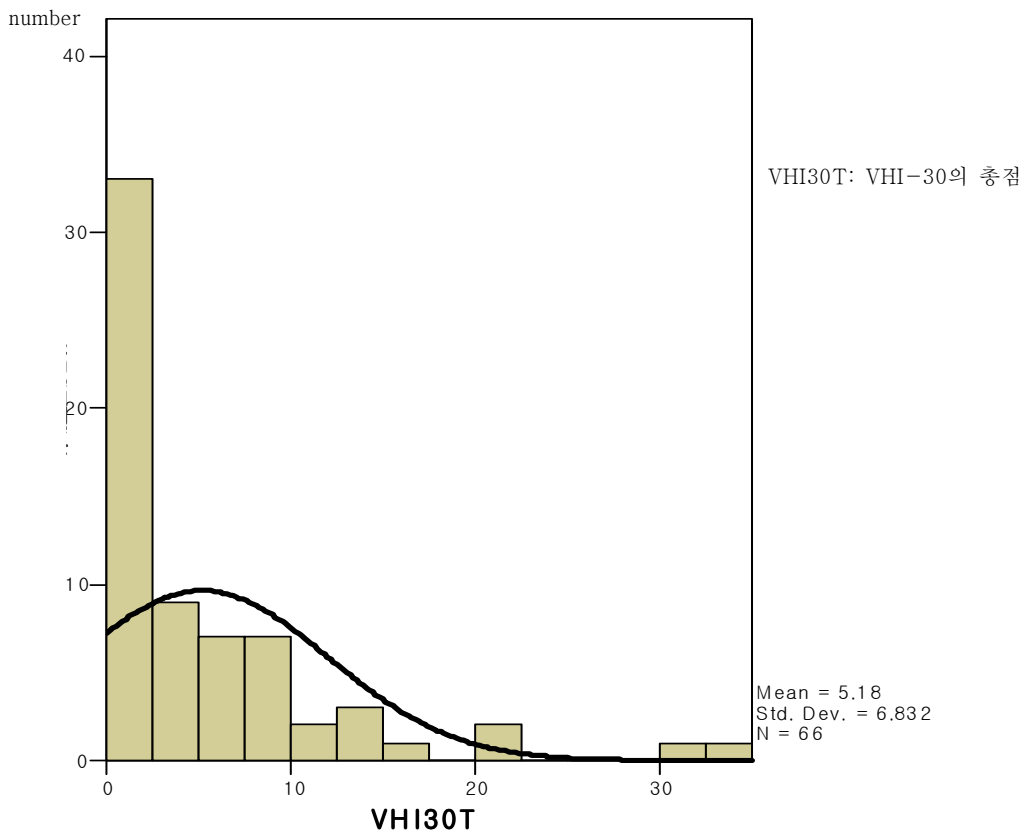


그림 2. 정상군의 VHI(Voice Handicap Index)-30의 총점 분포.

나. 신뢰도

(1) 검사-재검사 신뢰도

피어슨 상관(Pearson correlation) 분석에서 피험자 내 일치도가 매우 높았다 ($r=0.971$, $p<0.01$)(그림 3).

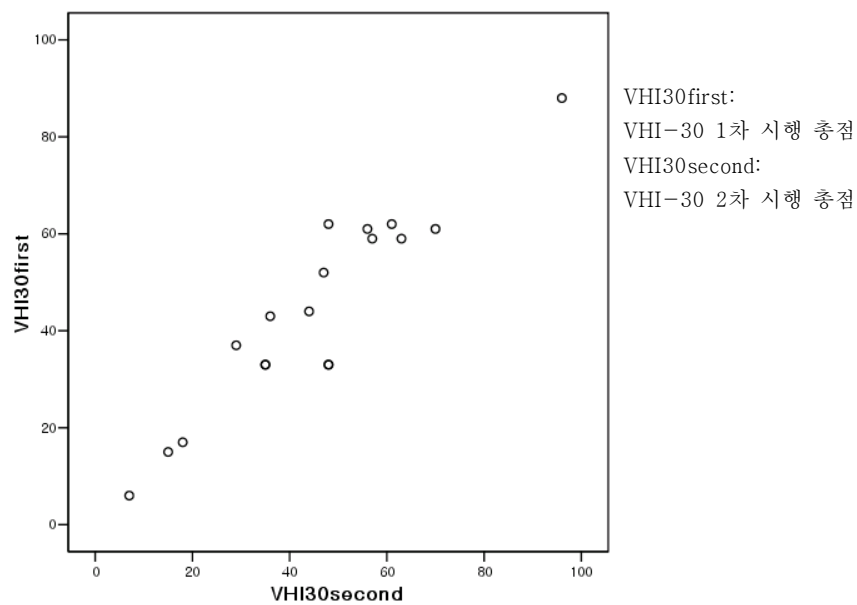


그림 3. VHI(Voice Handicap Index)-30의 검사-재검사 일치도.

(2) 문항내적일관성

환자군 156명을 대상으로 30문항 전체에 대한 Cronbach's α 값이 0.969로 매우 높은 신뢰도를 나타냈다.

(3) 문항변별도

환자군을 대상으로 각 문항의 점수와 총점 간의 상관성을 Spearman correlation 분석으로 알아본 결과, 모든 문항들이 총점과 상관성이 높았다($r=0.548\sim 0.846$, $p<0.01$)(표 4). 문항변별도가 높을수록 검사지의 신뢰도가 높다고 할 수 있다.⁵⁵ Ebel이 제시한 문항변별도 지수에 관한 평가기준에 따른 해석으로, 0.40 이상인 경우에 변별력이 높은 문항, 0.30 이상~0.40 미만인 경우에 변별력이 있는 문항, 0.30 미만의 경우에는 변별력이 낮은 문항으로 간주된다.⁵⁵

표 4. VHI(Voice Handicap Index)-30의 각 문항과 총점 간의 상관계수

문항	문항-총점 상관계수
F1	0.774*
P2	0.646*
F3	0.781*
P4	0.548*
F5	0.764*
F6	0.683*
E7	0.846*
F8	0.738*
E9	0.781*
P10	0.731*
F11	0.768*
F12	0.789*
P13	0.745*
P14	0.789*
E15	0.726*
F16	0.797*
P17	0.610*
P18	0.583*
F19	0.777*
P20	0.819*
P21	0.499*
F22	0.568*
E23	0.719*
E24	0.786*
E25	0.804*
P26	0.781*
E27	0.801*
E28	0.751*
E29	0.752*
E30	0.680*

* p < 0.01

다. 타당도

(1) 음성장애군과 정상군의 변별성

Mann Whitney 분석 결과, 환자군 156명과 정상군 66명을 대상으로 30문항의 총점 분포를 비교한 결과, 정상군에 비해 환자군에서 VHI-30 총점이 유의하게 높았다 ($z=-10.936$, $p<0.001$)(그림 4).

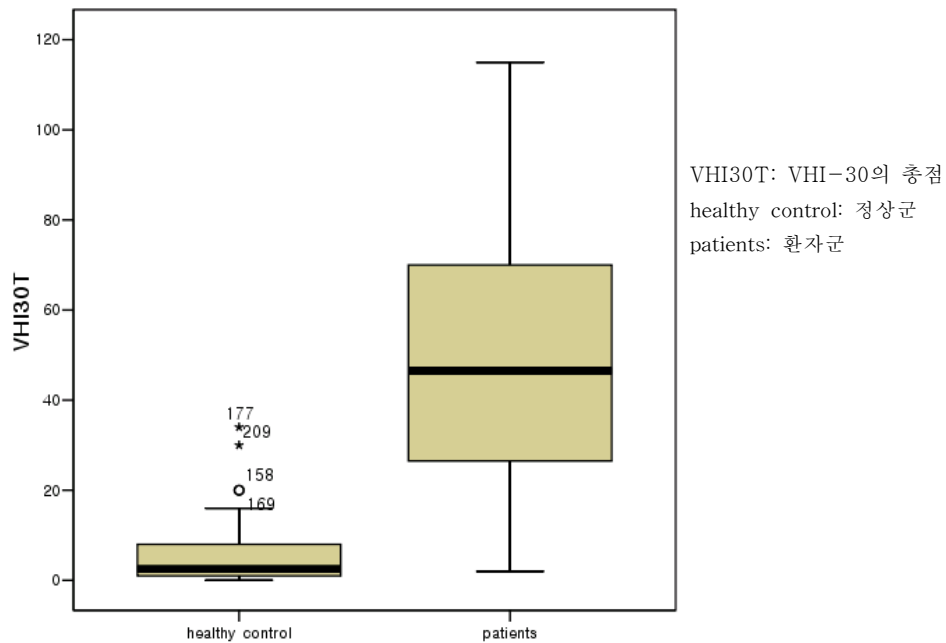


그림 4. 환자군과 정상군의 VHI(Voice Handicap Index)-30 총점 비교.

(2) 공인타당도

환자군을 대상으로 VHI-30의 총점과 IPVI의 1번 문항 점수를 Spearman correlation으로 분석하여 타당도를 확인한 결과, 비교적 높은 상관성을 보였다($r=0.664$, $p<0.01$). 환자군 및 정상군 모두를 대상으로 하여 VHI-30의 총점과 IPVI 1번 문항을 비교한 결과에서는 매우 높은 상관성을 보였다($r=0.845$, $p<0.01$)(그림 5).

타당도 지수에 의하여 평가하는 절대적 기준은 없으나, 성태제⁵⁵의 상관계수의 언어적 표현 기준에 의하면, 상관계수가 0.00~0.20이면 ‘타당도가 거의 없다’, 0.20~0.40이면 ‘타당도가 낮다’, 0.40~0.60이면 ‘타당도가 있다’, 0.60~0.80이면 ‘타당도가 높다’, 0.80~1.00이면 ‘타당도가 매우 높다’로 설명할 수 있다. 또한, 일반적으로 공인타당도에서 0.8 이상인 경우, 타당한 검사로 간주할 수 있다.⁵⁵

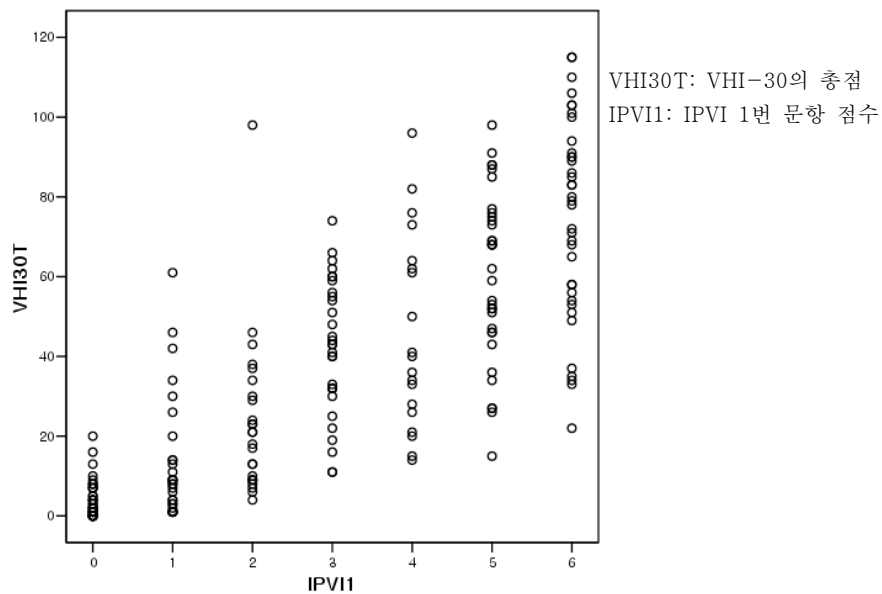


그림 5. 환자군 및 정상군의 VHI(Voice Handicap Index)-30 총점과 IPVI(Iowa Patient's Voice Index) 1번 문항 점수 간의 상관성.

(3) 요인분석

환자군을 대상으로 VHI- 30에 대한 요인분석을 시행한 결과, 원본과 마찬가지로 3개의 하위항목이 설정되었으나, 원본에 비해 하위항목 문항 구성 내용에 차이가 있었다. 첫째 항목은 총 13개로, F가 7개(F5, F12, F1, F3, F11, F8, F6), E가 4개(E15, E27, E9, E7), P가 2개(P2, P26)로 구성되어 기능적 요소가 중점적으로 포함된다고 할 수 있다(Eigen value: 16.151). 둘째 항목은 총 8개로, E가 5개(E29, E30, E24, E28, E25), F가 3개(F19, F22, F16)로 구성되어 비교적 정서적 요소를 나타낸다고 할 수 있다(Eigen value:1.792). 셋째 항목은 총 9개로, P가 8개(P13, P10, P4, P14, P18, P20, P21, P17), E가 1개(E23)로 구성되어 신체적 요소를 뚜렷하게 나타낸다고 할 수 있다(Eigen value: 1.611)(표 5, 그림 6).

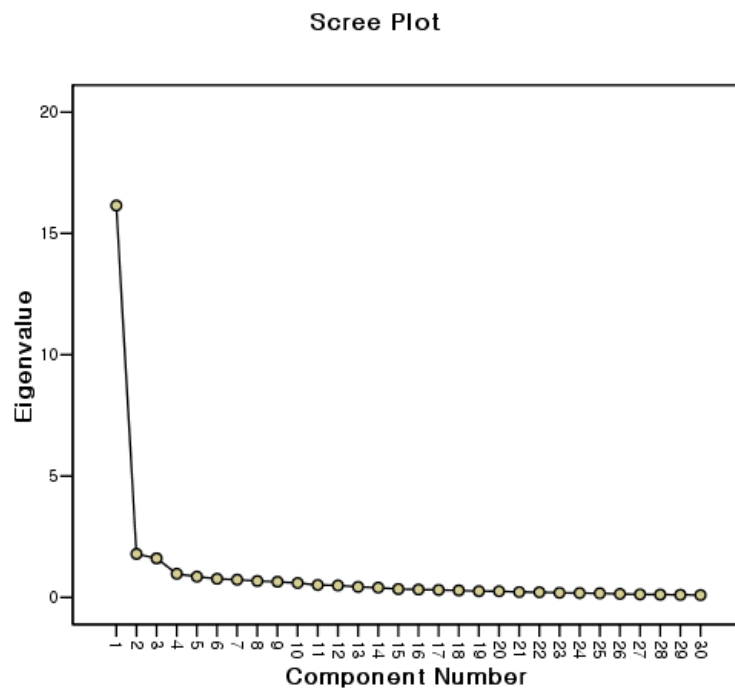


그림 6. VHI(Voice Handicap Index)-30의 요인 분석에 따른 scree plot.

표 5. VHI(Voice Handicap Index)-30의 요인분석 결과표

문항	구 성		
	1	2	3
F5	0.847		
F12	0.794		
F1	0.765		
F3	0.749		
E15	0.627		
F11	0.619		
E27	0.603		
P2	0.590		
E9	0.586		
F8	0.583		
F6	0.575		
E7	0.513		
P26	0.483		
E29		0.815	
E30		0.802	
E24		0.695	
E28		0.644	
E25		0.620	
F19		0.617	
F22		0.602	
F16		0.490	
P13			0.782
P10			0.708
P4			0.672
E23			0.660
P14			0.655
P18			0.584
P20			0.572
P21			0.567
P17			0.505

Rotated Component Matrix

2. VHI-30과 SF-36의 상관성

가. SF-36 결과 분석

SF-36은 수치가 증가할수록 건강이 상대적으로 좋지 않은 것으로 해석하였다. 따라서, 문항 중 3~19, 24, 25, 28, 29, 31~33, 35번은 나머지 문항과 척도의 방향이 달라 통계 분석을 실시하기 전에 대칭되는 값으로 수정하였다.

총 8개의 하위 항목은 다음과 같다. 신체적 기능은 3~12번에 해당되고, 신체적 건강 문제로 인한 역할 제한은 13~16번, 신체의 고통(통증)은 21~22번, 사회적 기능은 20, 32번, 일반적인 정신 건강은 24~26, 28, 30번, 정서적인 문제로 인한 역할 제한은 17~19번, 생동감, 원기, 또는 피로감은 23, 27, 29, 31번, 일반적 건강에 대한 의식정도는 1, 2, 33~36번에 해당되었다.

환자군 156명과 정상군 66명에 대한 총점을 비교한 결과, 두 군 모두 정규분포를 따랐고, 환자군에서 평균 78점(표준편차:18.4)을(그림 7), 정상군에서 평균 67.2점(표준편차: 13.6)을 보였다(그림 8). 두 군을 t 검정으로 비교한 결과, 정상군에 비해 환자군에서 유의하게 총점이 높았다($t(220) = 4.82, p < 0.001$)(그림 9).

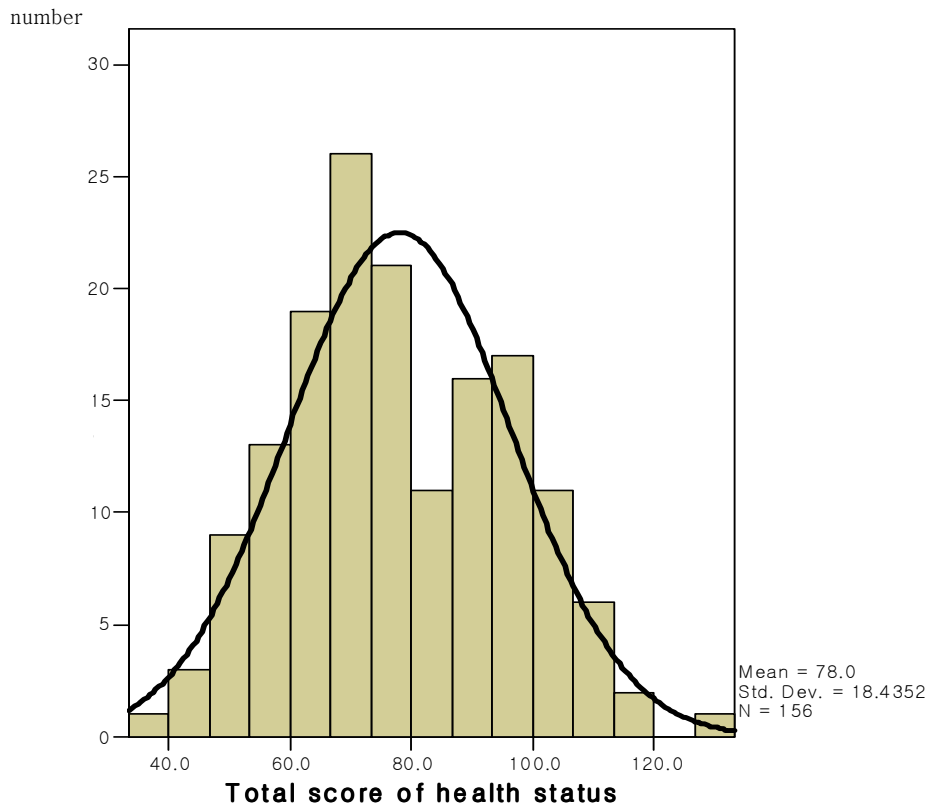


그림7. 환자군의 SF(Short Form)-36 총점 분포.

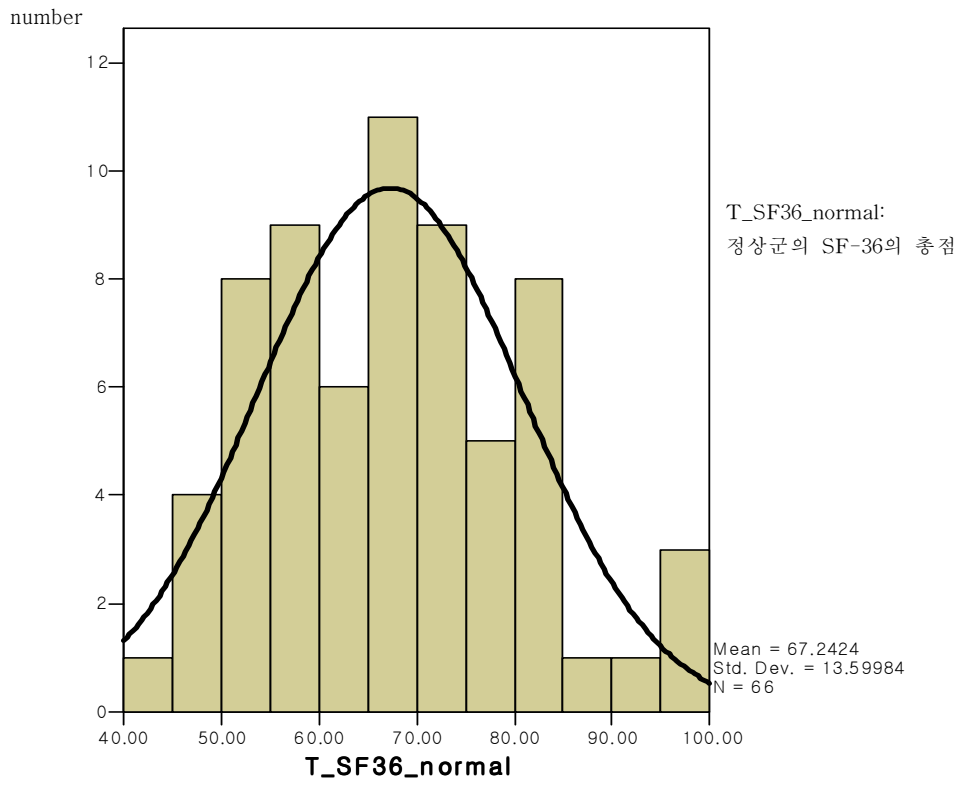


그림 8. 정상군의 SF(Short Form)-36 총점 분포.

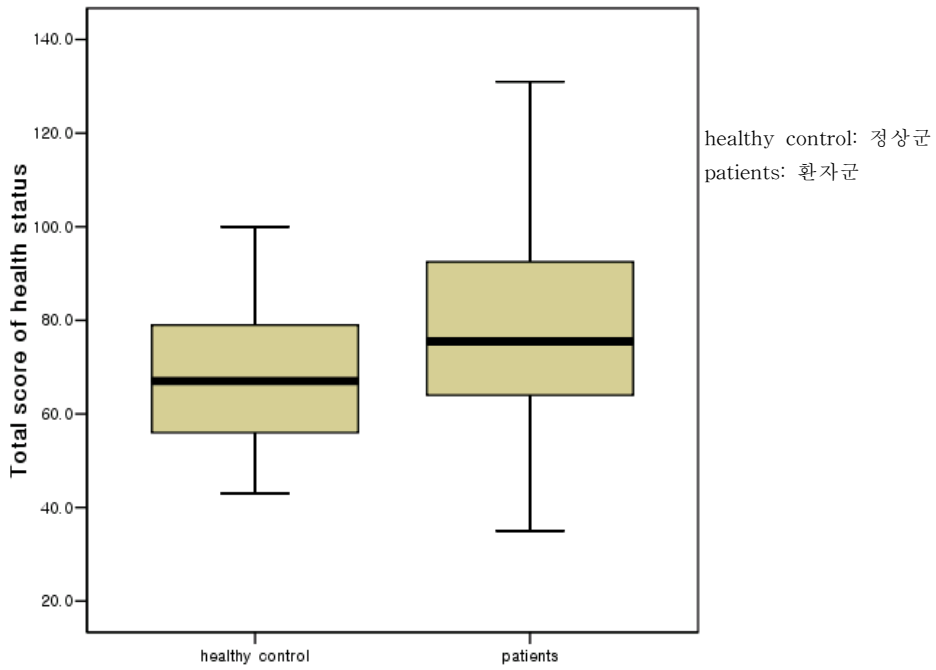


그림 9. 환자군과 정상군의 SF(Short Form)-36 총점 비교.

나. 상관분석

(1) VHI와 SF-36 간의 총점 비교

환자군을 대상으로 VHI-30 총점과 SF-36의 총점을 Spearman correlation으로 비교한 결과, 상관성이 있었다($r=0.372$, $p<0.01$) 환자군 및 정상군을 동시에 대상으로 비교한 결과에서는 상관성이 약간 더 증가하였다($r=0.409$, $p<0.01$)(그림 10).

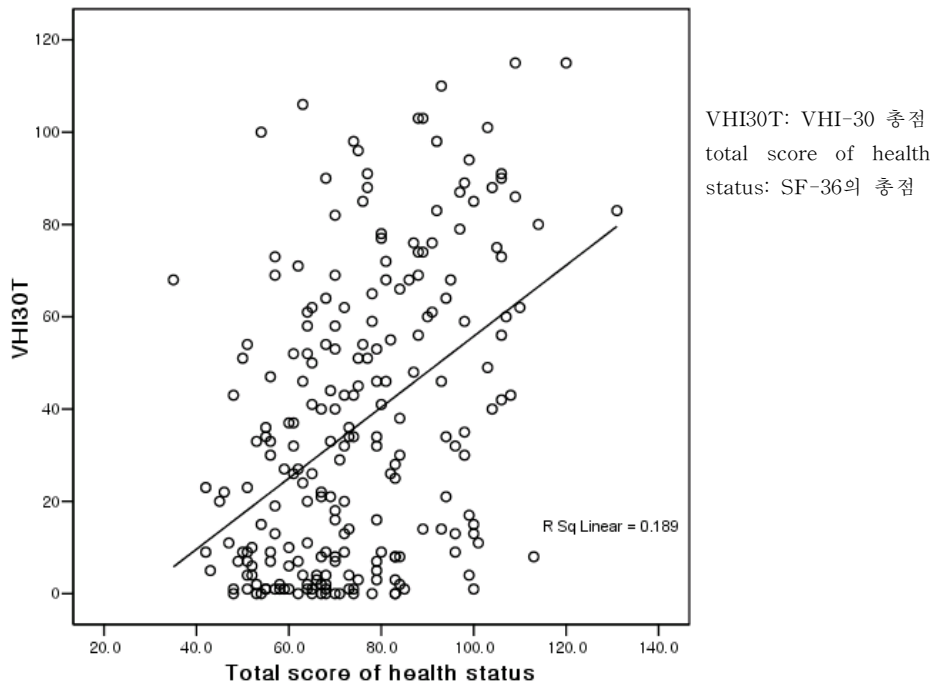


그림 10. VHI(Voice Handicap Index)-30과 SF(Short Form)-36의 총점 간 상관성.

(2) VHI 및 SF-36 간의 하위 항목 비교

환자군을 대상으로 VHI-30의 세 가지 하위 항목에 대해 SF-36의 여덟 가지 항목을 비교한 결과, 상관계수는 낮았으나 모든 항목 간에 유의한 상관성을 보였다. 그 중에서 SF-36의 사회 기능면의 항목이 다른 항목에 비해 VHI-30의 총점과 상대적으로 약간 더 높은 상관성을 보였다(표 6).

표 6. VHI(Voice handicap Index)-30 및 SF(Short Form)-36의 하위 항목 간 상관계수

	<i>Voice Handicap Index</i>		
	기능적 요소	신체적 요소	정서적 요소
<i>Short Form-36</i>			
신체적 기능	0.221*	0.252*	0.239*
신체적 건강 문제로 인한 역할 제한	0.263*	0.237*	0.263*
신체의 고통(통증)	0.284*	0.254*	0.298*
사회적 기능	0.415*	0.337*	0.416*
일반적인 정신건강	0.216*	0.209*	0.306*
정서적 문제로 인한 역할 제한	0.242*	0.198**	0.273*
생동감 또는 피로감	0.213*	0.296*	0.271*
일반적 건강에 대한 의식정도	0.167**	0.270*	0.234*

*p<0.01, **p<0.05

3. 질환별 VHI 및 SF-36 결과 비교

가. 질환 분류

환자군을 크게 네 가지, 신경학적 장애, 기능적 장애, 구조적 장애, 염증성 장애로 나누었다.¹ 신경학적 장애군은 총 21명으로 성대 마비, 경련성 발성장애가 해당되었다. 기능적 장애군은 총 80명으로 성대 폴립, 성대 결절, 라인케씨부종, 근긴장성 발성장애가 해당되었다. 구조적 장애군은 총 20명으로, 성대구증, 성대근위축, 후두유두종, 성대낭종, 피열연골골절, 후두격막이 해당되었다. 염증성 장애군은 총 35명으로 만성후두염, 후두인두역류염, 위산역류에 동반된 접촉성 궤양 및 육아종이 해당되었다(표 2).

나. 질환별 VHI 결과 비교

네 개의 질환군과 정상군을 대상으로 VHI-30 총점에 대한 차이를 Kruskal-Wallis test로 분석한 결과, 그룹 간 유의한 차이를 보였다($\chi^2(4) = 136.73$, $p < 0.001$). 보수적 통계의 Sheffé 사후검정 상, 정상군은 모든 환자군에 대해 유의하게 낮은 점수를 보였고, 환자군 중 신경학적 장애군은 나머지 기능적 장애군, 구조적 장애군 및 염증성 장애군에 대해 모두 유의하게 높은 점수를 보였다(표 7, 그림 11).

다. 질환별 SF-36 결과 비교

네 개의 질환별 환자군과 정상군을 대상으로 SF-36 총점에 대한 차이를 ANOVA로 분석한 결과, 그룹 간 유의한 차이를 보였다($F(4, 217) = 7.316$, $p < 0.001$). Sheffé 사후검정 상 신경학적 장애군은 구조적 장애군 및 정상군에 대해 유의한 차이를 보였다(표 8, 그림 12).

표 7. 질환별 환자군과 정상군의 VHI(Voice Handicap Index)-30 총점의 평균

	인원	평균±표준편차	최소치	최대치
환자군				
신경학적 장애	21	80.8±19.798	50	115
기능적 장애	80	44.7±25.214	7	101
구조적 장애	20	55.2±27.064	3	110
염증성 장애	35	39.0±27.699	2	100
정상군	66	5.2±6.832	0	34
총 계	222			

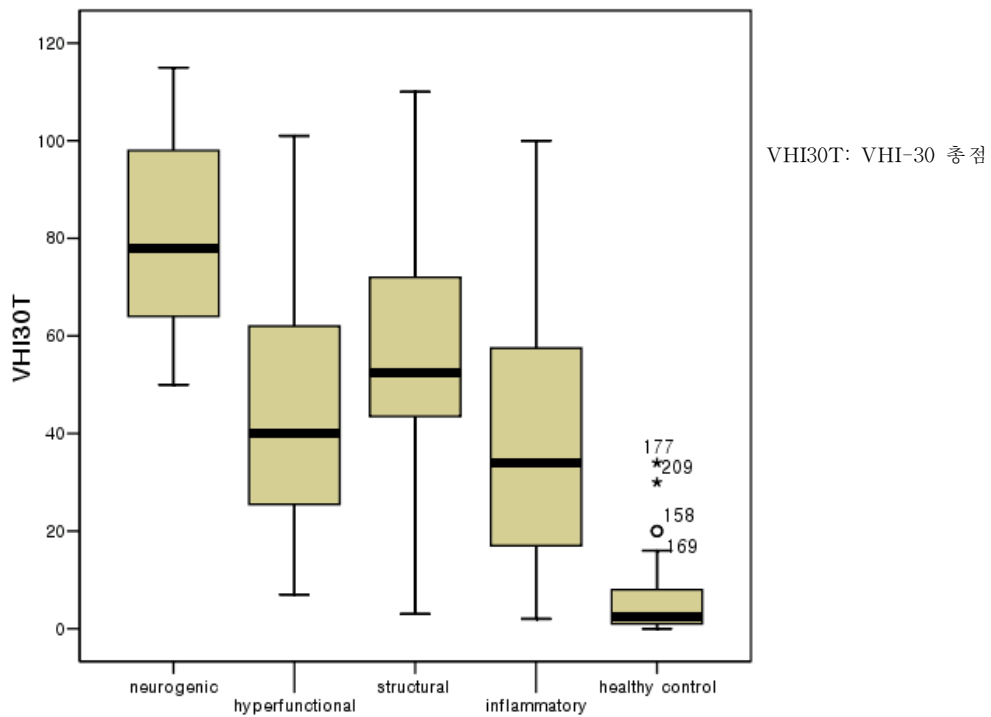


그림 11. 질환별 환자군과 정상군의 VHI(Voice Handicap Index)-30 총점 비교.

VHI30T: VHI-30 총점 neurogenic:신경학적 장애군
 hyperfunctional: 기능적 장애군 structural: 구조적 장애군
 inflammatory: 염증성 장애군 healthy control: 정상군

표 8. 질환별 환자군과 정상군의 SF(Short Form)-36 총점의 평균

	명수	평균±표준편차	최소치	최대치
환자군				
신경학적 장애	21	88.4±18.949	63	131
기능적 장애	80	77.6±17.655	35	110
구조적 장애	20	73.3±17.372	50	105
염증성 장애	35	75.4±18.928	42	114
정상군	66	67.2±13.600	43	100
총 계	222			

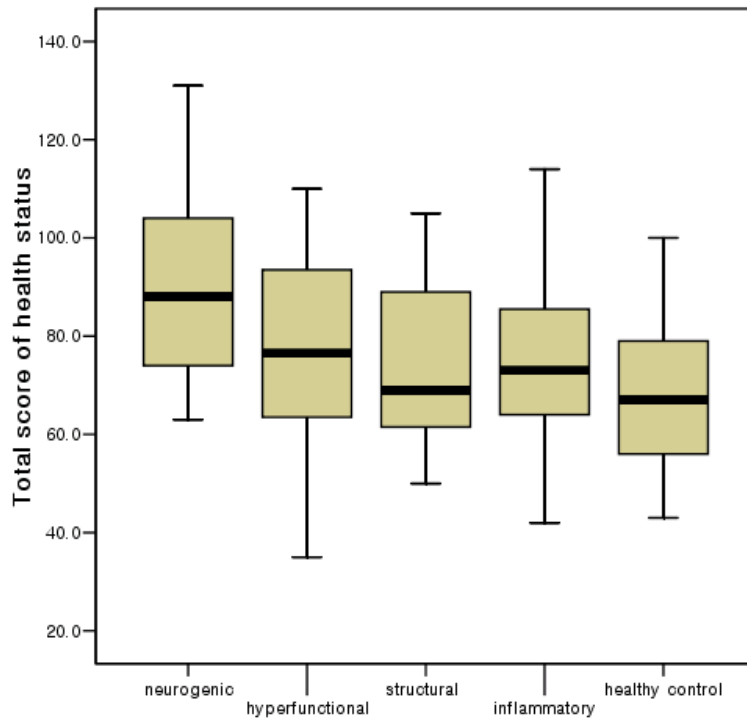


그림 12. 질환별 환자군과 정상군의 SF(Short Form)-36 총점 비교.

neurogenic:신경학적 장애군 hyperfunctional: 기능적 장애군
 structural: 구조적 장애군 inflammatory: 염증성 장애군
 healthy control: 정상군

4. VHI-30과 음향학적 결과 비교

가. 음향학적 결과 분석

환자군을 대상으로 MDVP를 시행한 결과, 성대마비 환자 1명에서 분석 불가를 보여 이를 제외한 155명에 대한 분석을 시행하였다. 결과는 표 9와 같으며, jitter(%)는 최소치 0.12%부터 최대치 15.27%를 보였고, 평균은 2.31%(중위수 값 1.61%)을 보였다. Shimmer(%)는 최소치 0.23%부터 최대치 24.12%를 보였고, 평균은 5.05%(중위수 값 4.17%)를 보였다. NHR은 최소치 0.063부터 최대치 21.69를 보였고, 평균은 0.39(중위수 값 0.14)를 보였다. 이들 결과 값은 수치가 증가할수록 발성의 안정도가 감소하는 것으로 해석한다.

표 9. 환자군의 MDVP(Multi Dimensional Voice Program) 파라미터 값

	평균±표준편차	중위수	최소치	최대치
jitter(%)	2.31±2.078	1.61	0.12	15.27
shimmer(%)	5.06±3.819	4.17	0.23	24.12
NHR ¹	0.40±1.933	0.14	0.06	21.69

¹ NHR: noise to harmonic ratio

나. VHI-30과 음향학적 결과 간의 상관성

VHI-30 총점에 대해 jitter는 낮은 상관성을 보였으나($r=0.209$, $p<0.001$), shimmer와 NHR은 상관성이 없었다(표 10). 음향학적 결과 내에서 jitter, shimmer, NHR 간에는 서로 상관성이 있는 것으로 나타났다.

표 10. VHI(Voice Handicap Index)-30 총점과 음향학적 파라미터 간의 상관계수

	VHI-30 총점	jitter	shimmer
jitter	0.209*		
shimmer	0.112	0.564*	
NHR ¹	0.058	0.489*	0.446*

* $p < 0.01$

¹ NHR: noise to harmonic ratio

5. VHI-30과 VHI-10 비교

가. VHI-10 신뢰도

(1) VHI-30에서 추출한 VHI-10와 VHI-10 간의 일치도

Spearman correlation으로 분석한 결과, 매우 높은 일치도를 보였다($r=0.916$, $p < 0.01$)(그림 13). 따라서 이후의 VHI-10에 대한 분석은 VHI-30에서 추출한 VHI-10 문항 결과로 시행하였다(이하 VHI-10과 동일하게 표기함).

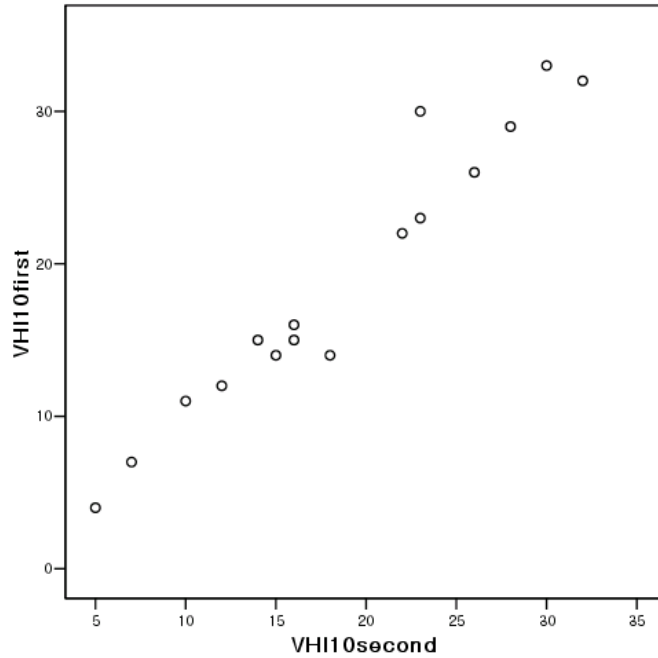


그림 13. VHI(Voice Handicap Index)-10의 일차 시행과 VHI-10 이차 시행 간의 일치도.

VHI10first: VHI-30에서 추출한 VHI-10의 1차 시행 총점
 VHI10second: VHI-10의 2차 시행 총점

(2) 문항내적일관성

VHI-30에서 추출한 VHI-10에 대한 Cronbach's α 는 0.916으로 매우 높은 문항내적일관성을 보였다.

나. VHI-10 타당도

(1) 환자군과 정상군의 총점 비교

Mann Whitney 분석에서, 환자군 156명과 정상군 66명을 대상으로 VHI-30에서 추출한 VHI-10문항의 총점에 대한 비교 결과, 정상군에 비해 환자군에서 유의하게 높았다($z=-10.744, p<0.001$)(그림 14).

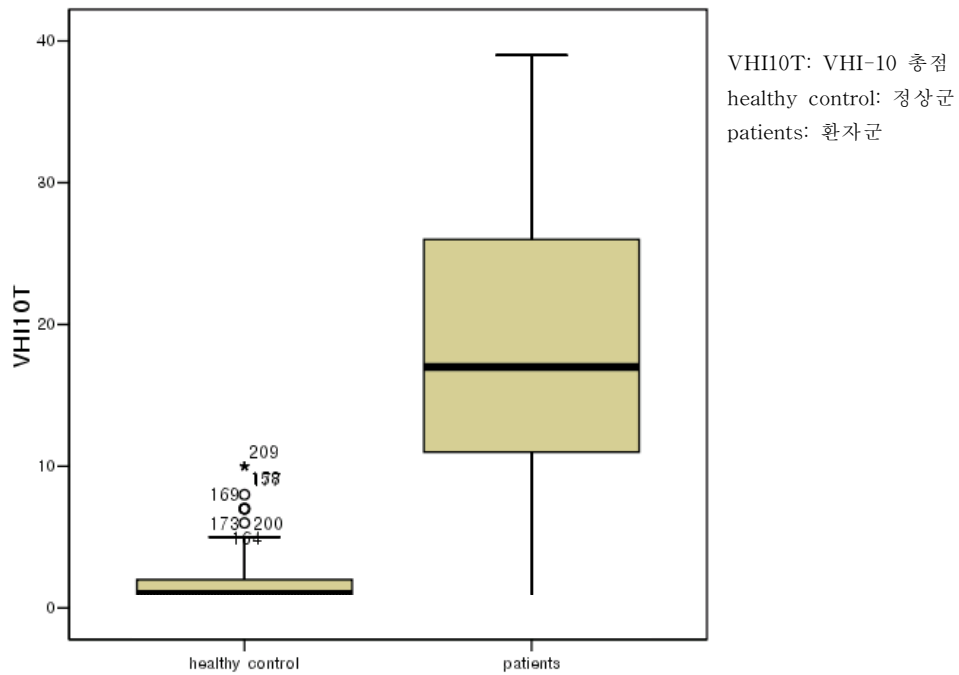


그림 14. 환자군과 정상군 간의 VHI(Voice Handicap Index)-10 총점 비교.

(2) 문항변별도

VHI-10의 각 문항들은 VHI-10의 총점과 상관성이 높았다($r=0.571\sim0.819, p<0.01$)(표 11).

표 11. VHI(Voice Handicap Index)-10의 각 문항 과 총점 간의 상관계수

문항	문항-총점 상관계수
F1	0.777*
F3	0.799*
P10	0.774*
P14	0.802*
F16	0.815*
P17	0.685*
F19	0.729*
F22	0.571*
E23	0.781*
E25	0.819*

* $p < 0.01$

(3) 공인타당도

환자군을 대상으로 VHI-30에서 추출한 VHI-10의 총점과 IPVI의 1번 문항을 Spearman correlation으로 분석한 결과, 높은 상관성을 보였다($r=0.681$, $p < 0.01$). 환자군과 정상군을 모두 대상으로 하여 VHI-10의 총점과 IPVI 1번 문항을 비교한 결과, 매우 높은 상관성을 보였다($r=0.848$, $p < 0.01$)(그림 15).

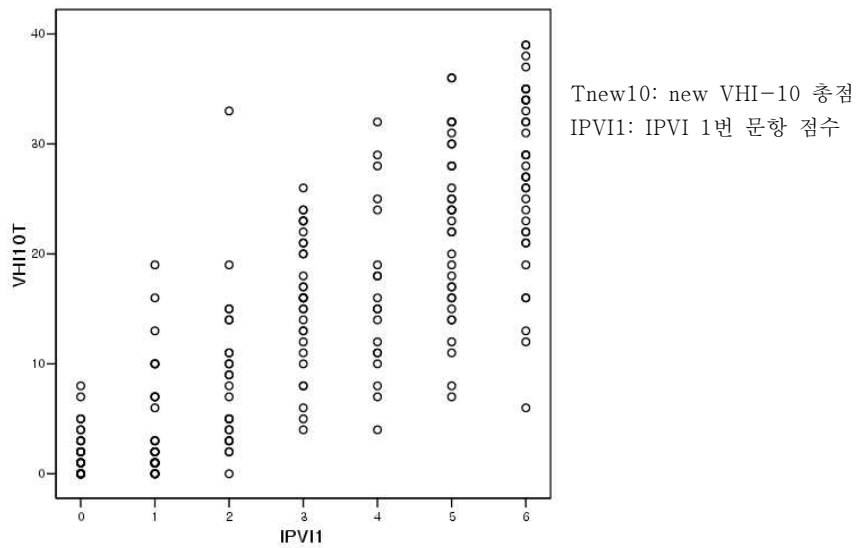


그림 15. VHI(Voice Handicap Index)-10 총점과 IPVI(Iowa Patient's Voice Index) 1번 문항 점수 간의 상관성.

(4) 요인분석

요인분석 상 하위항목으로 구분되지 않는 단일항목을 확인하였다.

(5) 문항 간 상관분석

환자군을 대상으로 VHI-10의 문항 간 상관성을 보기 위해 피어슨 상관(Pearson correlation) 분석하였다. 상관성이 높을수록 문항 간 내용에 유사성을 보이고, 상관성이 낮을수록 문항 내용에 독립성을 보인다고 해석할 수 있다. 상관성이 매우 높은 경우에는 둘 중 한 문항은 삭제가 요구될 수 있다. 분석 결과, 모든 문항 간에 유의한 상관관계를 보였다($p < 0.01$)(표 12). 상관 계수가 0.8 이상인 경우는 F1/F3이었고, 0.7이상~0.8 미만인 경우는 F16/E25이 해당되었다.

표 12. VHI(Voice Handicap Index)-10 문항 간 상관계수

	F1	F3	P10	P14	F16	P17	F19	F22	F23
F3	0.811*								
P10	0.539*	0.514*							
P14	0.630*	0.633*	0.666*						
F16	0.581*	0.554*	0.549*	0.588*					
P17	0.470*	0.497*	0.522*	0.514*	0.457*				
F19	0.529*	0.586*	0.440*	0.503*	0.597*	0.409*			
F22	0.302*	0.343*	0.380*	0.366*	0.452*	0.268*	0.388*		
E23	0.502*	0.537*	0.591*	0.575*	0.664*	0.508*	0.465*	0.413*	
E25	0.568*	0.581*	0.494*	0.601*	0.711*	0.509*	0.630*	0.425*	0.642*

* $p < 0.01$

다. VHI-30과 VHI-10의 상관성

환자군을 대상으로 VHI-30과 VHI-10 총점에 대한 상관 계수는 0.968로 높은 상관성을 보였다($p < 0.01$) (그림 16).

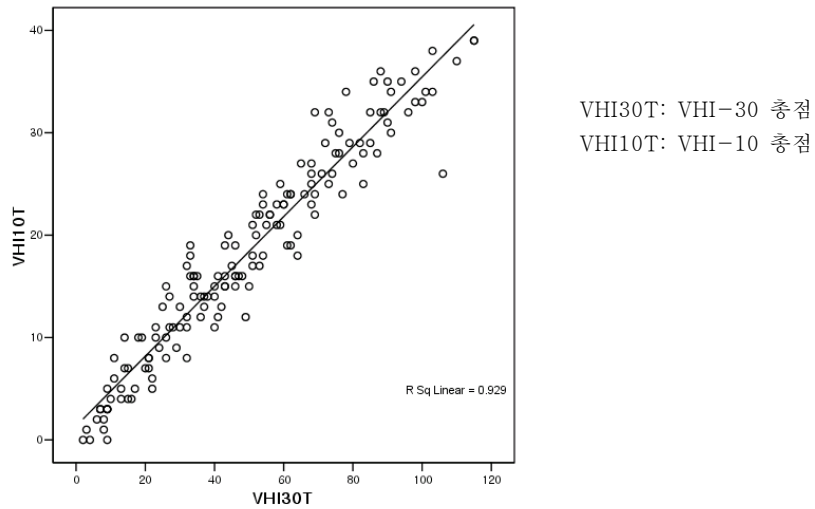


그림 16. VHI(Voice Handicap Index)-30 총점과 VHI-10 총점 간의 상관성.

라. VHI-30에 대한 총점 비율

기존 논문에서는 VHI-10 평균 총점 값 대 VHI-30 평균 총점 값 비율이 0.333보다 높은 경우에 VHI-10의 타당성이 상대적으로 높음을 증명하기도 하였다.^{7,27} 본 자료에서는 환자 전체에 대해서는 0.368이었고, 신경학적 장애 0.349, 기능적 장애 0.380, 구조적 장애 0.367, 염증성 장애 0.364로서, 모두 0.333에 비해 약간 높은 값을 보였다. 정상대조군의 경우에는 0.353을 보였다.

6. 타당성이 큰 10문항 선별

Rosen 등이 제시한 기존의 VHI-10 문항이 다른 언어 및 문화권에서도 마찬가지로 타당도가 높은 문항으로 선별될 수 있는지 확인해보고자 하였다.

가. 문항 분석

(1) 문항별 환자군-정상군 간 평균차 비교

VHI-30을 대상으로, 환자군과 정상군 간에 문항별 평균차를 확인하였다. 기존 논문에서 환자군과 정상군 간에 문항별 평균값의 차이가 클수록 두 군을 구별하는데 용이한 검사 문항임을 지적한 바 있다.⁷ 특히 평균차가 2.0이상 되는 문항은 음성장애와 정상을 뚜렷이 구분하는 문항이 된다고 하였다.⁷ 평균차가 큰 순서로는 P13(2.55), E23(2.35), P17(2.27), P10(2.23), P14(2.0), P4(1.91), P20(1.73), E25(1.71), E24(1.67), P18(1.64), P26(1.59), P21(1.58), E7(1.56), F16(1.53) 등이었다(표 13).

표 13. VHI(Voice Handicap Index)-30에 대한 환자군-정상군의 평균 차

문항	환자군-정상군의 평균 차
F1	1.35
P2	1.28
F3	1.29
P4	1.91
F5	1.06
F6	1.52
E7	1.56
F8	1.02
E9	1.12
P10	2.23
F11	1.46
F12	0.96
P13	2.55
P14	2.00
E15	1.40
F16	1.53
P17	2.27
P18	1.64
F19	0.90
P20	1.73
P21	1.58
F22	0.89
E23	2.35
E24	1.67
E25	1.71
P26	1.59
E27	1.03
E28	0.75
E29	1.14
E30	0.84

(2) 문항변별도

각 문항이 총점과 상관성이 높을수록 문항변별성이 있다고 볼 수 있으므로,⁴² 표 4 및 표 13에서 제시한 VHI 30문항에 대한 각 문항의 상관 계수가 높은 순서를 확인하였다. 상대적으로 높은 순서로는 E7(0.846), P20(0.819), E25(0.804), E27(0.801), F16(0.797), P14(0.789), F12(0.789) 등이었다.

(3) 문항별 Cronbach's α

문항 삭제 시 전체 Cronbach's α 가 높아지는 문항을 확인하였다. 문항 간 Cronbach's α 값의 차이가 거의 없는 편으로, 반드시 삭제를 요하는 문항은 없었다. 수치에 차이는 크지 않지만, 삭제 시 상대적으로 전체 Cronbach's α 를 증가시킨 문항은 0.970에 해당하는 문항 P21과, 0.969에 해당하는 문항 P2, P4, P17, P18이었다(표 14).

표 14. VHI(Voice Handicap Index)-30에 대한 문항-총점 상관성과 삭제 시 Cronbach's α

문항	문항-총점 상관계수	삭제 시 Cronbach's α
F1	0.774*	0.968
P2	0.646*	0.969
F3	0.781*	0.968
P4	0.548*	0.969
F5	0.764*	0.968
F6	0.683*	0.968
E7	0.846*	0.967
F8	0.738*	0.968
E9	0.781*	0.968
P10	0.731*	0.968
F11	0.768*	0.968
F12	0.789*	0.968
P13	0.745*	0.968
P14	0.789*	0.968
E15	0.726*	0.968
F16	0.797*	0.968
P17	0.610*	0.969
P18	0.583*	0.969
F19	0.777*	0.968
P20	0.819*	0.967
P21	0.499*	0.970
F22	0.568*	0.969
E23	0.719*	0.968
E24	0.786*	0.968
E25	0.804*	0.968
P26	0.781*	0.968
E27	0.801*	0.968
E28	0.751*	0.968
E29	0.752*	0.968
E30	0.680*	0.968

*p<0.01

나. 새로 선택된 10문항(new VHI-10)

문항별 평균차가 큰 문항 순서를 우선하여 나열한 후, 문항별 점수와 총점 간의 상관성이 비교적 높은 문항을 순차적으로 확인하고, 문항 삭제 시 전체 Cronbach's α 가 약간 높아지는 P4, P17, P18, P21을 제거한 후, 그 결과 10개의 문항을 추출하였다. 10개로 추출된 문항(new VHI-10, 이하 nVHI-10)은 P13, E23, P10, P14, P20, E25, E24, P26, E7, F16이었다(부록 6). 이에 대한 타당도를 기존 VHI-10과 비교해 보았다.

(1) VHI-30과 nVHI-10의 상관성

환자군을 대상으로 VHI-30의 총점과 nVHI-10의 총점을 Spearman correlation으로 분석한 결과, 높은 상관성을 보였다($r=0.961$, $p<0.01$)(그림 17).

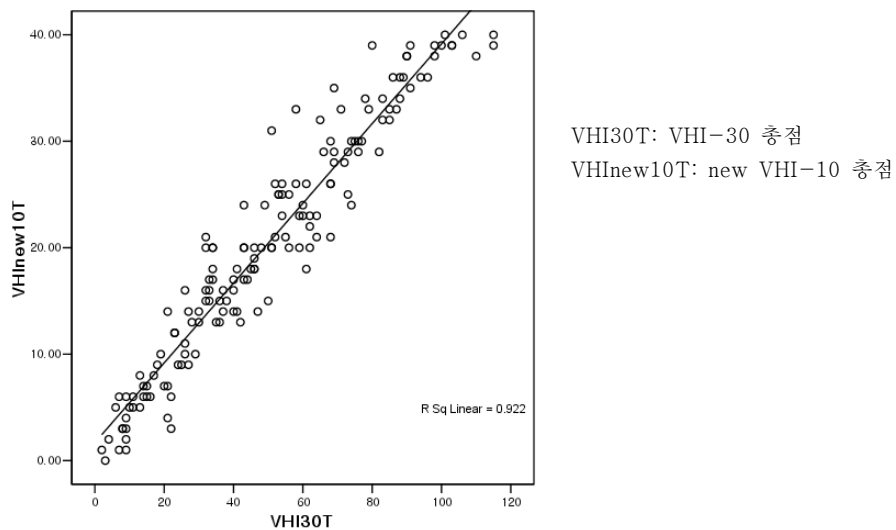


그림 17. VHI(Voice Handicap Index)-30의 총점과 nVHI(new Voice Handicap Index)-10의 총점 간의 상관성.

(2) VHI-30에 대한 총점 비율

nVHI-10의 평균 총점 값 대 VHI-30 평균 총점 값 비율은 환자 전체에 대해서는 0.408, 신경학적 장애 0.376, 기능적 장애 0.417, 구조적 장애 0.429, 염증성 장애 0.411로서, 0.333이상의 값을 보였고, 기존 VHI-10의 0.368에 비해 높은 수치를 보였다. 정상대조군의 경우에는 0.262를 보였다.

(3) 공인타당도

환자군을 대상으로 nVHI-10의 총점과 IPVI의 1번 문항을 Spearman correlation으로 분석한 결과, 비교적 높은 상관성을 보였다($r=0.707$, $p<0.01$). 환자군과 정상군을 모두 대상으로 하여 nVHI-10의 총점과 IPVI 1번 문항을 비교한 결과, 보다 높은 상관성을 보였다($r=0.868$, $p<0.01$)(그림 18). 기존 VHI-10에 비해 상관 계수 값이 모두 약간씩 높아졌다.

(4) 요인분석

기존 VHI-10과 마찬가지로 요인분석 상 nVHI-10에서는 하위항목이 나누어지지 않았다.

(5) 문항 간 상관분석

환자군을 대상으로 nVHI-10의 문항 간 상관성을 보기 위해 피어슨 상관분석을 한 결과, 모든 문항 간에 유의한 상관관계를 보였다($p<0.01$)(표 15). 상관 계수가 0.8 이상인 경우는 없었고, 0.7이상~0.8 미만인 경우는 E24/E25, P14/P20, E7/P20, P10/P13, F16/E25, P13/P14가 해당되었다.

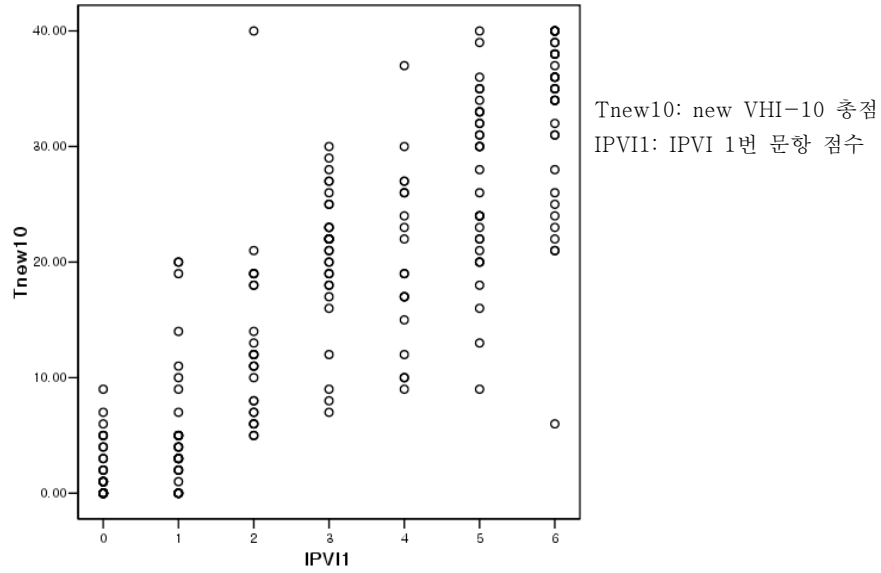


그림 18. nVHI(new Voice Handicap Index)-10 총점과 IPVI(Iowa Patient's Voice Index) 1번 문항 점수 간의 상관성.

표 15. nVHI(new Voice Handicap Index)-10 문항 간 상관계수

	P13	E23	P10	P14	P20	E25	E24	P26	E7
E23	0.574*								
P10	0.714*	0.591*							
P14	0.702*	0.575*	0.666*						
P20	0.631*	0.579*	0.604*	0.774*					
E25	0.543*	0.642*	0.494*	0.601*	0.690*				
E24	0.557*	0.637*	0.521*	0.517*	0.662*	0.778*			
P26	0.579*	0.583*	0.558*	0.642*	0.662*	0.657*	0.620*		
E7	0.608*	0.590*	0.591*	0.673*	0.732*	0.614*	0.637*	0.641*	
F16	0.562*	0.664*	0.549*	0.588*	0.641*	0.711*	0.653*	0.577*	0.677*

*p<0.01

7. ‘노래 부르기’ 문항 분석

가. 문항 분석

(1) 문항별 환자군-정상군 간 평균 차 비교

‘노래 부르기’ 문항이 포함된 VHI-33을 대상으로, 환자군과 정상군 간에 문항별 평균차를 확인하였다. 평균차가 큰 순서로는 S31(2.55), P13(2.55), E23(2.35), P17(2.27), P10(2.23), P14(2.0), P4(1.91), S32(1.82), P20(1.73), E25(1.71), E24(1.67), P18(1.64) 등이었다(표 16).

표 16. VHI(Voice Handicap Index)-30 및 ‘노래 부르기’ 세 문항에 대한 환자군-정상군의 평균 차

문항	환자군-정상군의 평균 차
F1	1.35
P2	1.28
F3	1.29
P4	1.91
F5	1.06
F6	1.52
E7	1.56
F8	1.02
E9	1.12
P10	2.23
F11	1.46
F12	0.96
P13	2.55
P14	2.00
E15	1.40
F16	1.53
P17	2.27
P18	1.64
F19	0.90
P20	1.73
P21	1.58
F22	0.89
E23	2.35
E24	1.67
E25	1.71
P26	1.59
E27	1.03
E28	0.75
E29	1.14
E30	0.84
S31	2.55
S32	1.82
S33	1.33

(2) 문항변별도

각 문항 점수와 총점 간의 상관성으로 문항변별도를 알아보았다. 모든 문항에서 상관 계수가 0.501부터 0.843로 비교적 높은 변별력을 보였다(표 17). 상관 계수가 상대적으로 높은 순서로는 E7(0.843), P20(0.815), E25(0.804), E24(0.797), E27(0.796), F16(0.790), P14(0.786) 등이었다.

(3) 문항별 Cronbach's α

문항 삭제 시 전체 Cronbach's α 가 높아지는 문항을 확인하였다. 13개 문항이 0.968, 16개 문항이 0.969, 4개의 문항이 0.970으로 문항별 차이는 거의 없었다. 이 중 0.970에 해당하는 문항은 P4, P21, P22, S33이었다(표 17).

표 17. VHI(Voice Handicap Index) 및 ‘노래 부르기’ 세 문항에 대한
문항-총점 상관성과 삭제 시 Cronbach's α

문항	문항-총점 상관계수	삭제 시 Cronbach's α
F1	0.764*	0.969
P2	0.633*	0.969
F3	0.768*	0.968
P4	0.550*	0.970
F5	0.750*	0.969
F6	0.679*	0.969
E7	0.843*	0.968
F8	0.736*	0.969
E9	0.763*	0.968
P10	0.737*	0.969
F11	0.772*	0.968
F12	0.772*	0.968
P13	0.754*	0.969
P14	0.786*	0.968
E15	0.701*	0.969
F16	0.790*	0.968
P17	0.604*	0.969
P18	0.588*	0.969
F19	0.776*	0.968
P20	0.815*	0.968
P21	0.501*	0.970
F22	0.561*	0.970
E23	0.724*	0.969
E24	0.797*	0.968
E25	0.804*	0.968
P26	0.778*	0.968
E27	0.796*	0.968
E28	0.749*	0.969
E29	0.754*	0.969
E30	0.681*	0.969
S31	0.659*	0.969
S32	0.629*	0.969
S33	0.577*	0.970

* $p < 0.01$

나. '노래 부르기' 가 포함된 10 문항(sVHI-10) 제작

문항별 평균차가 큰 문항 순서를 우선하여 나열한 후, 문항별 점수와 총점 간의 상관성이 비교적 높은 문항을 순차적으로 확인하고, 문항 삭제 시 전체 Cronbach's α 가 약간 높아지는 P4, P21, P22, S33를 제거한 후, 그 결과 10개의 문항(singing VHI-10, 이하 sVHI-10)을 선별하였다. sVHI-10은 S31(2.55), P13(2.55), E23(2.35), P17(2.27), P10(2.23), P14(2.0), S32(1.82), P20(1.73), E25(1.71), E24(1.67)이었고(부록 7), 이에 대한 타당도를 기존 VHI-10과 비교해 보았다.

(1) VHI-30에 대한 총점 비율

sVHI-10의 평균 총점 값 대 VHI-30 평균 총점 값 비율이 환자 전체에 대해서는 0.456, 신경학적 장애 0.389, 기능적 장애 0.488, 구조적 장애 0.457, 염증성 장애 0.475로, 0.333보다 높은 수치를 보였으며, 기존 VHI-10에 비해 비율이 높았다. 정상대조군의 경우 0.374를 보였다.

(2) 공인타당도

환자군을 대상으로 sVHI-10의 총점과 IPVI의 1번 문항의 총점을 Spearman correlation으로 분석한 결과, 비교적 높은 상관성을 보였다($r=0.690$, $p<0.01$). 환자군과 정상군을 모두 대상으로 하여 sVHI-10의 총점과 IPVI의 1번 문항의 총점을 비교한 결과, 보다 높은 상관성을 보였다($r=0.884$, $p<0.01$)(그림 19). 기존 VHI-10에 비해 상관 계수 값이 모두 약간씩 높아졌다.

(3) 요인분석

기존 VHI-10과 마찬가지로 요인분석 상 sVHI-10에서는 하위 항목이 나누어지

지 않았다.

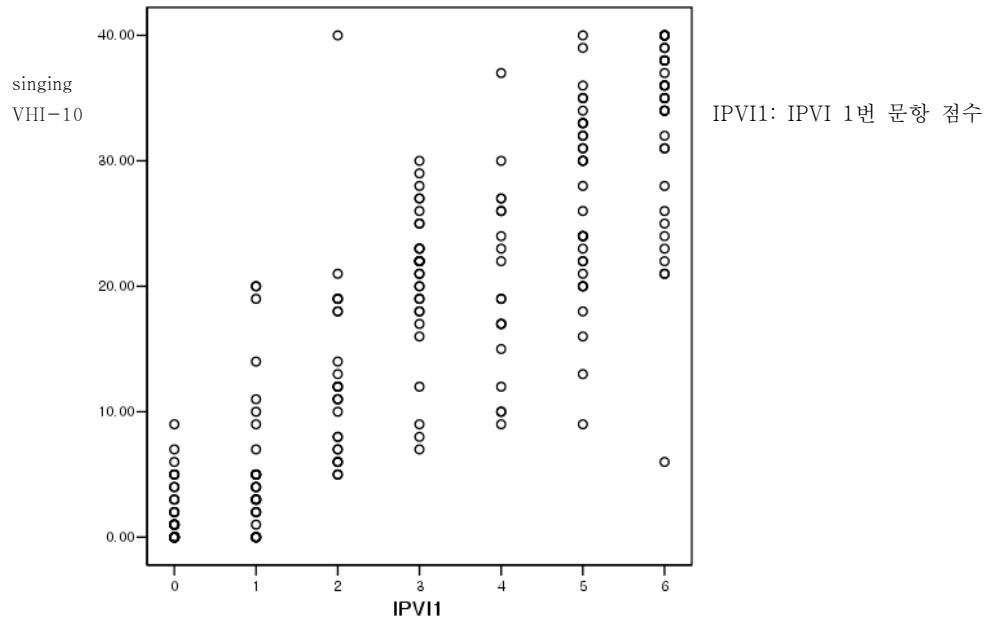


그림 19. sVHI(singing Voice Handicap Index)-10 총점과 IPVI(Iowa Patient's Voice Index) 1번 문항 점수 간의 상관성.

(4) 문항 간 상관분석

환자군을 대상으로 sVHI-10의 문항 간 상관성을 보기 위해 피어슨 상관분석을 한 결과, 모든 문항 간에 유의한 상관관계를 보였다($p < 0.01$)(표 18). 상관계수가 0.8 이상인 경우는 없었고, 0.7이상~0.8 미만인 경우는 E24/E25, P14/P20, P13/P10, P13/P14가 해당되었다.

표 18. sVHI(singing Voice Handicap Index)-10 문항 간 상관계수

	S31	P13	E23	P17	P10	P14	S32	P20	E25
P13	0.594*								
E23	0.585*	0.574*							
P17	0.436*	0.545*	0.508*						
P10	0.569*	0.714*	0.591*	0.522*					
P14	0.531*	0.702*	0.575*	0.514*	0.666*				
S32	0.613*	0.509*	0.441*	0.336*	0.464*	0.487*			
P20	0.506*	0.631*	0.579*	0.390*	0.604*	0.774*	0.515*		
E25	0.456*	0.543*	0.642*	0.509*	0.494*	0.601*	0.490*	0.690*	
E24	0.460*	0.557*	0.637*	0.435*	0.521*	0.517*	0.578*	0.662*	0.778*

*p<0.01

IV. 고 찰

한국어판 VHI의 유용성

본 연구에서 확인된 한국어판 VHI-30은 내용타당도를 거쳐 신뢰도 및 타당도 면에서 유의한 측정도구임을 밝혔다. Jacobson 등(1997)이 VHI를 제작하였을 때는 피험자 내 일치도가 $r=0.92$, 문항 전체에 대한 Cronbach's α 는 0.95이었고, 공인타당도를 대신한 4점 척도의 음성 문제 정도와 비교했을 때 VHI 총점과 높은 상관성($r=0.60$)을 보였다.⁸ 본 연구와 Jacobson 등의 연구를 비교해보면 대부분 근사치의 결과를 보였거나, 본 연구에서 약간 더 높은 결과 값을 보였다.

Jacobson 등(1997)의 연구에서는 요인 분석을 하지 않고, 세 가지 하위영역으로 임의의 구분을 하였다. 세 가지 하위영역에 대한 유효성 검증은 이후 여러 연구에서 시행되어 왔지만, 일관성 있는 결과를 보이지 않았다. Rosen 등(2004)은 하위항목으로 나누어 분석하는 것에 의의가 없음을 지적한 반면,⁷ Lam 등(2006)은 세 가지 하위영역으로 구분이 가능하였다고 보고한 바 있었다.⁴² 본 연구에서는 Jacobson 등이 제시한 하위영역의 내용과는 일치하지는 않았지만, 요인분석 상 비교적 유사한 내용 구성의 세 가지 하위영역으로 구분이 가능함을 확인하였다. 특히 신체적인 요인은 Jacobson의 구분과 비교해보았을 때 가장 유사한 내용 구성을 보였다. 이는 신체적인 요인이 후두의 불편감이나 음성 자체의 문제에 국한하기 때문에 비교적 일차적인 문제를 나타내는 특징을 보이기 때문이라고 생각한다. 반면, 기능적 요인이나 정서적 요인은 생리적인 문제로 야기되는 이차적인 사회 심리적인 면을 나타내는 것으로 신체적인 요인에 비해 구별이 뚜렷하지 않았지만, 삶의 질을 다양한 관점에서 평가할 수 있게 하는 역할을 한다. 따라서 VHI-30이 고유의 세 가지 항목으로 구분이 되는 것에 유효성이 있다고 볼 수 있으므로, 향후 좀 더 안정된 요인분석을 위해 대상자 수를 늘여 이를 지지하는 후속 연구가 요구된다.

Cohen 등(2006)은 1966년부터 2003년까지 VHI와 SF-36의 이전 연구들의 자

료들을 종합하여 비교한 결과, 음성장애 환자가 정상군에 비해 VHI와 SF-36에서 모두 유의하게 나쁜 점수를 얻는 결과를 보였고, 특히 신경학적 질환이 기타의 염증성 질환이나 과다기능에 의한 질환 등에 비해 유의하게 나쁜 점수를 얻었다고 보고하였다.³⁸ Benninger 등(1998)은 SF-36의 하위항목 중에서 특히 사회적 기능 항목이 VHI의 총점과 하위항목들과 상관성이 가장 높았고, 성대마비가 다른 기능적 음성장애나 기타 종양, 부종 등의 질환에 의한 음성장애에 비해 치료 전에 모두 VHI와 SF-36에서 더 나쁜 점수를 받았다고 하였다.³⁶ 본 연구에서도 성대마비 및 경련성발성장애에 해당하는 신경학적 장애군이 기타 기능적, 구조적, 염증성 장애군에 비해 VHI와 SF-36에서 모두 유의하게 증가되어 음성장애로 인한 삶의 질이 상대적으로 낮음을 나타내었다. 이를 통해 VHI가 장애 그룹 간 비교 연구에 사용 가능함을 시사하였다. 또한 VHI와 SF-36의 하위항목을 비교한 결과에서도 선행연구와 같은 결과로서, 사회적 기능 항목이 다른 항목에 비해 상관성이 높았다. 본 연구에서는 위와 같이 선행연구의 결과와 유사 또는 일치하는 결과를 보인 바, 대상자 자료수집이 비교적 적절하였음을 간접적으로 확인할 수 있었다.

VHI의 총점과 MDVP의 각 파라미터 간의 상관성을 알아보았을 때, jitter에서는 상관성이 있었으나, shimmer와 NHR에서는 상관성이 없었다. 그러나 jitter 경우에도 상관 계수 값이 높지 않았으므로, 평가자의 객관적 검사인 음향학적 평가와 환자의 주관적인 평가 간에는 상관성이 적은 것으로 해석할 수 있다. 이는 환자의 주관적인 평가는 평가자의 객관적인 검사와는 독립적인 평가임을 시사하며, 결국 환자마다 주관적으로 느끼는 음성 문제의 심각도는 객관적인 평가만으로는 충분하지 않고 개인의 기준에 따른 주관적인 평가를 시행할 필요가 있음을 재확인하게 하였다.

문항수를 단축한 VHI

본 연구에서 확인된 한국어판 VHI-10는 VHI-30과 마찬가지로 신뢰도 및 타당도 면에서 유의한 측정도구임을 밝혔다. VHI-10에 대한 선행 연구들과 비교해보면, 본 연구 결과와 유사함을 보였다. 히브리어 번안본⁴³에서는 전체 Cronbach's α 는 0.949

였고, VHI-30과 VHI-10간의 총점에 대한 상관 계수는 0.970($p < 0.01$)이었다. 중국어 번안본⁴²에서는 전체 Cronbach's α 는 0.95였고, VHI-30과 VHI-10간의 상관 계수는 0.970($p < 0.01$)이었다. 또한 공인타당도를 대신하여 음성장애에 대한 7점 척도의 자가 측정 결과와 비교하였을 때 높은 상관성을 보였다($r = 0.726, p < .001$).

그러나 Rosen 등(2004)의 주장⁷과는 달리, 본 자료에서는 VHI-30에 대한 하위 항목 구분이 가능하였기에, VHI-10이 VHI-30에 비해 유용성이 더 있다고 볼 수는 없었다. 따라서 연구의 목적에 따라, 하위항목을 비교할 경우에는 VHI-30을 사용하는 것이 바람직하겠고, 하위항목 구분 없이 설문 시간 단축을 위한 경우에는 VHI-10를 사용하는 것이 바람직하겠다.

Rosen 등(2004)의 연구에서는 VHI-10 평균 총점 값 대 VHI-30 평균 총점 값 비율을 질환별로 알아보았고, 최소 0.366부터 최대 0.474를 보이는 등 평균 0.4 이상을 보여, VHI-10의 뚜렷한 유용성을 주장하였다.⁷ Lam 등(2006)의 연구에서는 신경학적 음성장애에서 0.362, 구조적 음성장애에서 0.380, 염증성 음성장애에서 0.349를 보여, Rosen 등의 연구에 비해 비율이 낮은 편이나, 0.333보다는 높은 점수를 유지하였다.⁴² 본 연구에서는 환자군 전체에 대해 0.368을 보여, Lam의 결과와 유사한 값을 보였다. Rosen 등의 연구에서는 타 연구에 비해 비율이 매우 높은 편인데, 이는 Rosen 등이 영어권 환자를 대상으로 VHI-30에서 VHI-10을 추린 기준을 중국어권 또는 한국어권의 환자에게 적용할 때는 그 차이가 있을 수 있음을 가정해볼 수 있다.

따라서 본 연구에서는 VHI 30문항에서 가장 타당한 10개의 문항을 추렸을 때, 기존 VHI-10의 문항과 어떠한 차이를 보이는 지 확인을 한 바, 기존 VHI-10에 해당하는 문항이 5개가 포함된 새로운 10개의 문항(nVHI-10)을 설정할 수 있었다(부록 6). nVHI-10에 해당되는 문항은 F23, P10, P14, E25, F16이고, 그 외에 추가된 문항은 P13, P20, E24, P26, E7이었다. 신체적인 항목이 5개, 기능적인 항목이 1개, 정서적인 항목이 4개로, 기존의 VHI-10에 비해 기능적인 항목에 대한 문항이 감소되고, 상대적으로 신체적인 항목이 뚜렷이 증가됨을 알 수 있었다. VHI-30에 대한 총점 값 비율이 기존 VHI-10의 0.368에 비해 0.408로 높은 비율을 보였으므로, 민감성이 상대적으로 더 높다고 설명할 수 있다. Rosen 등(2004)은 선정 기준을 치료

전후 편차를 많이 보인 문항, 음성장애인과 정상 대조군 간에 차이를 많이 보인 문항, 그리고 전문가의 종합적 의견에 의한 추천이었으나, 본 연구에서는 음성장애인과 정상 대조군 간에 차이를 많이 보인 문항, 문항변별도가 높은 문항, 삭제 시 Cronbach's α 가 증가하는 문항을 기준하였으므로, 향후 선정 기준의 차이에 대한 결과 해석을 고려해야 한다. F22 문항은 Rosen의 연구에서도 변별성이 크지 않았지만, 전문가에 의한 내용 타당도 추천으로 포함이 된 경우인데, 한국인을 대상으로 했을 때는 오히려 삭제를 추천할 수 있다. 설문문의 내용이 수입 감소 여부를 확인하는 것이므로 경제활동을 하는 대상자가 아닌 경우에는 이에 대한 응답이 어려울 수 있고(따라서, '일을 한다면'의 가정 문구를 문장 앞에서 첨부하는 것이 요구될 수 있음), 판매 서비스 직종 외로는 음성 문제로 인한 수입의 감소를 직접적으로 느끼지 못하는 편이기 때문이다. 또한 서구사회에 비해 수입으로 자신의 역량을 평가하는 사회정서가 많지 않는 것도 한 이유가 된다고 생각한다.

본 연구에서는 말 수준의 음성장애 정도를 알아보는 VHI 30문항뿐만 아니라, 우리나라의 문화적 특성을 고려한 '노래 부르기' 문항을 첨부하여 전체 문항에 대한 분석을 실시하였다. 추려진 10개의 문항은 기존 VHI-10 중 5개(P14, E25, P10, E23, P17)가 해당 되었고, P20, P13, E24와 '노래 부르기' 두 문항(S31, S32)이 첨가되었다(부록 7). sVHI-10의 평균 총점 값 대 VHI-30 평균 총점 값 비율이 0.456으로, 기존 VHI-10의 0.368과 nVHI-10의 0.408보다 비율이 높았다. 따라서 한국인들에게 '노래 부르기'는 일상생활에서 나타나는 음성 기능 중 상대적으로 매우 중요한 부분을 차지한다고 유추할 수 있다.

sVHI-10은 nVHI-10의 문항과 구성에 유사성을 보여, 7개의 문항이 중복되었다. 따라서 sVHI-10에서도 nVHI-10과 마찬가지로 신체적인 항목이 두드러진 구성을 보였다. S31을 포함한다면 신체적인 문항은 총 6개이고, 본 연구의 요인 분석 결과에 기준하여 E23까지 포함한다면 총 7개의 문항이 신체적인 요인임을 알 수 있었다. 본 연구에서는 환자군과 정상군 간 점수 차이를 많이 보인 문항을 일차 선별 기준으로 하였기에, 이러한 현상은 환자군에서 기능적 또는 정서적인 항목보다는 신체적인 항목에 해당하는 문항의 점수가 높았음을 그 원인으로 해석할 수가 있다. 따라서 문항의 수가 적을 경우에는 일차적으로 생리적 불편감을 표현한 신체적인 문항이 두드러짐

을 알 수 있었다.

본 연구의 제한점 및 향후 과제

언어적 특성 상 일대일의 완전한 번역은 가능하지 않을 수 있으나, 번안본은 최대한 원문에 가까워야 국제적인 공통 연구에도 사용할 수 있다. 그 반면, 언어에 대한 문화적 차이를 고려해야 모국어 사용자들에게 문항 변별력을 높일 수 있을 것이다. 본 연구에서 제시한 번안본 또한 앞으로 사용해가면서 더 적절한 번안이 될 수도 있을 것이다.

본 연구에서 제시한 VHI-30 및 VHI-10에 대한 결과는 선행 연구들과 비교하여 신뢰도 및 타당도가 있음을 확인하였다. 그러나, 이는 고전검사이론⁵⁵에 근거한 분석이었으므로 실험에 참여하는 대상군이 달라질 때마다 결과가 다를 수 있는 단점이 있다. 만약 문항반응이론⁵⁵에 의한 분석을 한다고 하여도, VHI와 같이 주관적 척도로 구성된 문항에 대해 정-오답을 분석하는 일반적인 문항분석은 실행하기 어려우므로 향후 이에 대한 수정된 모형을 제시해야한다고 본다.

본 연구에서는 시행하지 않았으나 앞으로는 음성 장애 특성에 대한 반응 패턴 분석이 요구된다. 문항 전체의 총점뿐만 아니라 개별 문항에 대한 특성을 살펴본다면 문항에 따라서는 환자의 진단명이나 사회활동에 따라 특정 문항에 대해서는 반응에 일정한 경향을 보이는 것이 있을 수 있다. 예를 들어, VHI의 신체적 문항 중, ‘말을 할 때 숨이 차다’의 경우, 성문폐쇄부전을 보이는 환자가 주로 호소할 수 있겠고, ‘목소리를 내려면 힘을 주어야 나오는 것 같다’의 경우, 내전형 경련성발성장애에서 뚜렷하게 호소하는 증상이 될 수 있다. VHI에는 포함되지 않지만, ‘자고난 후 목소리 변화가 뚜렷하다’라는 문항은 위산역류로 인한 후두질환의 경우에 변별성이 높을 수 있다.⁵⁶ 이렇듯 문항마다 대상자의 요소에 따라 특성을 보이는 기준을 제시한다면, 차후에는 전체 문항 점수뿐만 아니라, 대상자에 따라 해당 문항을 선택하여 개별 점수를 부여하거나, 차후에 일정한 가산점을 주는 것을 고려할 수도 있을 것이다. 또한 직업적인 특성을 고려하여 전문적인 음성 사용자를 위한 새로운 문항 개발이 요구될 수도 있다.

임상에서 사용 간편성을 위해 문항 수는 줄이면서, 효과적으로 측정을 하는 문항 제작은 앞으로도 진행될 전망이다. 본 연구에서는 VHI-10을 기준한 민감도를 알아보기 위해 문항 수 10개를 기준하여 nVHI-10과 sVHI-10을 비교하였다. 그러나 nVHI-10 또는 sVHI-10의 임상적 유용성을 확인하기 위해서는 향후 이에 대한 신뢰도 및 타당도에 대한 평가가 추가적으로 시행되어야 할 것이다. 만약 10개의 문항보다 문항 수는 줄고 측정의 효과성은 유지된다면, 이를 적극 제안할 수 있을 것이다. 예를 들어, P14(‘목소리를 내려면 힘을 주어야 나오는 것 같다’)와 P20(‘말할 때는 애를 많이 쓰게 된다’)의 문항 간 상관성은 0.714 정도로 다른 문항들에 비해 비교적 상관성이 높은 문항에 속하는 바, 전문가에 의해 의미 상 큰 차이가 없는 경우로 인정이 된다면, 둘 중 하나의 문항만을 채택하는 것도 가능할 것이다.

본 연구에서는 현재까지 이용도가 가장 높은 VHI를 대상으로 신뢰도와 타당도를 보았지만, 기타 음성장애에 관한 다른 설문지에 대한 비교 연구도 계속 필요하겠다. 특히 공인타당도를 보기 위해서 신뢰도 및 타당도를 확인한 다양한 설문지도 마련되어야 하겠다.

V. 결 론

한국어로 번안하여 제시한 VHI-30, VHI-10는 신뢰도 및 타당도가 있음을 확인하였다. VHI-30은 세 가지의 하위항목으로 분석이 가능하였고, 질환군별 유의한 차이를 보인 바, 이들을 기준한 비교 연구에 음성장애의 삶의 질을 측정하는 도구로서 사용 가능함을 보였다. 한편, 임상에서 사용에 간편성을 위해서는 VHI-30 대신, VHI-10을 사용하는 것에 유용성을 보였다.

VHI와 MDVP의 음향학적 지수 간에는 상관성이 없거나 낮았으므로, 환자의 주관적 음성 문제 인식 정도와 객관적 평가에 의한 음성 문제 정도에는 그 기준에 차이가 있음을 확인하였다. 따라서 두 평가가 독립적인 평가임을 제시한다.

본 연구에서 문항 분석을 통해 VHI-30 문항 중 10개의 문항을 새로 선택한 nVHI-10과 한국의 문화적 특성을 반영한 ‘노래 부르기’ 문항이 포함된 sVHI-10은 모두 기존 VHI-10에 비해 상대적으로 민감도가 높음을 확인하였다. 따라서 향후 문항 선정을 할 경우에는 언어 및 문화권의 차이에 대한 고려를 하는 것이 바람직하리라 생각한다.

VHI는 음성장애-심리평가도구로서 비단 조사 연구의 이용 가치뿐만 아니라 임상에서도 매우 유용하게 사용될 수 있다. 이를 실시하는 동안 환자가 스스로 문제의 정도를 점수화하면서 그 과정에서 자신의 음성 문제에 대한 동기가 증가되고, 언어치료사 입장에서도 제공된 치료에 대한 전후 호전도 평가에 사용이 가능하며, 환자의 요구 수준에 맞는 치료 방향 결정에도 도움이 되어 이에 대한 이용도는 앞으로도 증가할 것이다.

참고문헌

- 1) Boone DR, McFarlane SC, Von Berg SL. The voice and voice therapy. 7th ed. Needham Heights: Allyn & Bacon; 2005.
- 2) Koschkee DL, Rammage L. Voice care in the medical setting. San Diego: Singular Publishing Group; 1997.
- 3) Kent RD, Ball MJ. Voice quality measurement. San Diego: Singular Publishing Group; 2000.
- 4) Kent RD, Read C. The acoustic analysis of speech. San Diego: Singular Publishing Group; 1992.
- 5) Hirano M, Bless DM. Videostroboscopic examination of the larynx. San Diego: Singular Publishing Group; 1993.
- 6) Dejonckere PH, Obbens C, de Moor GM, Wieneke GH. Perceptual evaluation of dysphonia: reliability and relevance. *Folia Phoniatr* 1993;45:76-83.
- 7) Rosen CA, Lee AS, Osborne J, Zullo T, Murry T. Development and validation of the Voice Handicap Index-10. *Laryngoscope* 2004;114:1549-56.
- 8) Jacobson BH, Johnson A, Grywalski C, Silbergleit A, Jacobson G, Benninger MS, et al. The Voice Handicap Index(VHI): development and validation. *Am J Speech Lang Pathol* 1997;3:66-70.
- 9) Hogikyan ND, Sethuraman Girish. Validation of an instrument to measure Voice-Related Quality of Life(V-RQOL). *J Voice* 1999;13: 557-69.
- 10) Deary IJ, Wilson JA, Carding PN, MacKenzie K. VoiSS: a patient-derived Voice Symptoms Scale. *J Psychosom Res* 2003;54:483-9.
- 11) Deary IJ, Webb A, MacKenzie K, Wilson J, Carding PN. Short, self-report voice symptom scales: psychometric characteristics of the Voice Handicap Index-10 and the Vocal Performance Questionnaire. *Otolaryngol Head Neck Surg* 2004;131: 232-5.

- 12) Bach KK, Belafsky PC, Wasylik K, Postma GN, Koufman JA. Validity and reliability of the Glottal Function Index. *Arch Otolaryngol Head Neck Surg* 2005;131:961-4.
- 13) World Health Organization. International classification of impairments, disabilities, and handicaps: a manual of classification relating to the consequences of disease. Geneva: World Health Organization; 1980.
- 14) Siupsinskiene N, Adamonis K, Toohill RJ. Quality of life in laryngopharyngeal reflux patients. *Laryngoscope* 2007;117:480-4.
- 15) Spector BC, Netterville JL, Billante C, Clary J, Reinisch L, Smith TL. Quality of life assessment in patients with unilateral vocal cord paralysis. *Otolaryngol Head Neck Surg* 2001;125:176-82.
- 16) Rosen CA, Murry T, Zinn A, Zullo T, Sonbolian M. Voice Handicap Index change following treatment of voice disorders. *J Voice* 2000;14:619-23.
- 17) Wingate JM, Ruddy BH, Lundy DS, Lehman J, Casiano R, Collins SP, et al. Voice Handicap Index results for older patients with adductor spasmodic dysphonia. *J Voice* 2005;19:124-31.
- 18) Fulljames N, Harris S. Voice outcome measures: correlations with patient's assessment of their condition and the effectiveness of voice therapy. *Logoped Phoniatr Vocol* 2006;31:23-5.
- 19) Spielman J, Ramig LO, Mahher L, Halpern A, Gavin WJ. Effects of an extended version of the Lee Silverman Voice Treatment on voice and speech in Parkinson's disease. *Am J Speech Lang Pathol* 2007;16:95-107.
- 20) Alves CB, Loughran S, MacGregor FB, Dey JI, Bowie LJ. Bioplastic medialization therapy improves the quality of life of terminally ill patients with vocal cord palsy. *Clin Otolaryngol* 2002;27:387-91.
- 21) Carding PN, Horsley IA, Docherty GD. Measuring the effectiveness of voice therapy in a group of forty-five patients with non-organic dysphonia. *J Voice* 1999;13:557-69.

- 22) Criteria for determining disability in speech-language disorders. Agency for Healthcare Research and Quality(AHRQ). Evidence Report/ Technology Assessment. January 2002;52.
- 23) Wilson JA, Webb PN, Carding IN, Steen IN, MacKenzie K, Deary IJ. The Voice Symptom Scale(VoiSS) and the Vocal Handicap Index(VHI): a comparison of structure and content. Clin Otolaryngol Allied Sci 2004;29:169-74.
- 24) Rosen CA, Murry T. Voice Handicap Index in singers. J Voice 1999;14:370-7.
- 25) Ma EP, Yiu EM. Voice activity and participation profile: assessing the impact of voice disorders on daily activities. J Speech Lang Hear Res 2001;44:511-24.
- 26) Yiu EM, Ma EP. Voice activity limitation and participation restriction in the teaching profession: the need for preventative care. J Med Speech-Lang Pathol 2002;10:51-60.
- 27) Hogikyan ND, Wodchic WP, Spak C, Kileny PR. Longitudinal effects of botulinum toxin injections on voice-related quality of life (V-RQOL) for patients with adductory spasmodic dysphonia. J Voice 2001;15:576-86.
- 28) Hogikyan ND, Wodchis WP, Terrell JE, Bradford CR, Esclamado RM. Voice-related quality of life(V-RQOL) following Type I thyroplasty for unilateral vocal fold paralysis. J Voice 2000;14:378-86.
- 29) Gliklich RE, Glovsky RM, Montgomery WM. Validation of a voice outcome survey for unilateral vocal cord paralysis. Otolaryngol-Head Neck Surg 1999;120:153-8.
- 30) Terrell J, Nanavait K, Esclamado R, Bishop J, Bradford C, Wolf G. Head and neck cancer-specific quality of life: instrument validation. Arch Otolaryngol Head Neck Surg 1997;123:1125-32.
- 31) Weymuller EA, Alsaeeaf R, Yueh B, Deleyiannis FW, Coltrera MD.

- Analysis of the performance characteristics of the University of Washington Quality of Life Instrument and its modification (UW-QOL-R). Arch Otolaryngol Head Neck Surg 2001;127:489-93.
- 32) Franic DM, Bramlett RE, Bothe AC. Psychometric evaluation of disease specific quality of life instruments in voice disorders. J Voice 2005;19:300-15.
- 33) Reed PJ, Moore D. SF-36 as a predictor of health states. Value Health 2000;3:202-7.
- 34) Krischke S, Weigelt S, Hoppe U, Kollner V, Klotz M, Eysholdt U, et al. Quality of life in dysphonic patients. J Voice 2005;19:132-7.
- 35) Wilson JA, Deary IJ, Millar A, Mackenzie K. The quality of life impact of dysphonia. Clin Otolaryngol 2002;27:179-82.
- 36) Benninger MS, Ahuja AS, Gardner G, Grywalski C. Assessing outcomes for dysphonia patients. J Voice 1998;12:540-50.
- 37) 최세준, 한주희, 박영준, 최승화, 김상윤, 남순열. 후두미세수술을 시행 받은 환자의 삶의 질. 대한음성언어의학회지 2005;16:146-51.
- 38) Cohen SM, Dupont WD, Courey MS. Quality-of-life impact of non-neoplastic voice disorders: a meta-analysis. Ann Otol Rhinol Laryngol 2006;115:128-34.
- 39) Dejonckere PH, Remacle M, Fresnel-Elbaz E, Woisard V, Crevier-Buchman L. Differentiated perceptual evaluation of pathological voice quality: reliability and correlations with acoustic measurements. Rev Larynol Otol Rhinol 1996;117:219-24.
- 40) 윤영선, 이은경, 백정환, 손영익. 성대 결절의 호전도 측정을 위한 청지각적 평가 및 화상회선경검사 분석 비교. 대한이비인후과학회지 2005;48:1501-5.
- 41) Fullijames N, Harris S. Voice outcome measure: correlations with patient's assessment of their condition and the effectiveness of voice therapy. Logoped Phoniatr Vocol 2006;31:23-35.

- 42) Lam PKY, Chan KM, Ho WK, Kwong E, Yiu EM, Wei WI. Cross-cultural adaptation and validation of the Chinese Voice Handicap Index-10. *Laryngoscope* 2006;116:1192-8.
- 43) Amir O, Tavor Y, Leibovitzh T, Ashkenazi O, Michael O, Primov-Fever A, et al. Evaluating the validity of the Voice Handicap Index-10 among Hebrew speakers. *Otolaryngol-Head Neck Surg* 2006;135:603-7.
- 44) Woisard V, Bodin S, Puech M. The Voice Handicap Index: impact of the translation in French on the validation. *Rev Laryngol Otol Rhinol* 2004;125:307-12.
- 45) Hakkesteegt MM, Wieringa MH, Gerritsma EJ, Feenstra L. Reproducibility of the Dutch version of the Voice Handicap Index. *Folia Phoniatr Logop* 2006;58:132-8.
- 46) Guimaraes I, Abberton E. An investigation of the Voice Handicap Index with speakers of Portuguese: preliminary data. *J Voice* 2004;18:71-82.
- 47) Hsiung MW, Lu P, Kang BH, Wang HW. Measurement and validation of the Voice Handicap Index in voice-disordered patients in Taiwan. *J Laryngol Otol* 2003;117:478-81.
- 48) Petersen MA, Groenvold M, Bjorner JB, Aaronson N, Conroy T, Cull A, et al. Use of differential item functioning analysis to assess the equivalence of translations of a questionnaire. *Qual Life Res* 2003;12:373-85.
- 49) 최두영, 최선명, 임채길, 남순열. 원 목소리 환자에서의 음성장애지수의 유용성. *대한이비인후과학회지* 2002;45:706-10.
- 50) 박성신. 초등학교 여교사를 대상으로 한 음성위생법 효과성. *이화여자대학교 석사학위논문*; 2004.
- 51) Kim A, Lim E. Effects of different types of practice in cross-cultural test adaptation of affective measures. *Korean J Psychol* 2003;22:89-113.
- 52) Choi Y, Mericle A, Harachi TW. Using Rasch analysis to test the cross-cultural item equivalence of the Harvard Trauma Questionnaire and

- the Hopkins Symptom Checklist across Vietnamese and Cambodian immigrant mother. J Appl Meas 2006;7:16-38.
- 53) Kay Elemetrics Corp. Operations manual: Multi-Dimensional Voice Program(MDVP). Pine Brook; 1993.
- 54) 손영익, 윤영선, 권중근, 추광철. 발성 시 음도 및 강도의 변화가 음성분석검사 결과에 미치는 영향. 대한음성언어학회지 1997;8:12-17.
- 55) 성태제. 문항제작 및 분석의 이론과 실제. 2판. 서울: 학지사; 2004.
- 56) Bove MJ, Rosen C. Diagnosis and management of laryngopharyngeal reflux disease. Curr Opin Otolaryngol Head Neck Surg 2006;14:116-23.

부록 1. IPVI(Iowa Patient's Voice Index) 중 1번 문항

다음 문항 중 본인에 해당하는 정도의 숫자에 동그라미 또는 V로 표시하십시오.

(0이면 정상을 나타내고, 숫자가 증가할수록 문제의 정도가 증가하여 6이 가장 심한 정도를 나타냄)

1. 내 목소리에

문제가 없다 0 1 2 3 4 5 6 문제가 심각하다

부록 2. VHI(Voice Handicap Index)-30

다음의 질문에 대해 본인이 느끼는 증상이 어느 정도인지 숫자에 동그라미(또는 V표)로 표시 하십시오.

- 0=전혀 그렇지 않다
- 1=거의 그렇지 않다
- 2=가끔 그렇다
- 3=자주 그렇다
- 4=항상 그렇다

- F1 목소리 때문에 상대방이 내 말을 알아듣기 힘들어한다 0 1 2 3 4
- P2 말을 할 때 숨이 차다
- F3 시끄러운 곳에서는 사람들이 내 말을 이해하기 어려워한다
- P4 하루 중에도 목소리가 자주 변한다
- F5 집안 어디서든 내가 부르는 말소리를 가족들이 잘 듣지 못 한다
- F6 목소리 때문에 전화통화를 가급적 줄인다
- E7 목소리 때문에 타인과 대화를 할 때 긴장을 한다
- F8 내 목소리 때문에 여러 사람이 모인 자리를 피하게된다
- E9 내 목소리 때문에 사람들은 짜증이 날 것이다
- P10 사람들이 나에게 목소리가 왜 그러냐고 묻는다
- F11 내 목소리 때문에 친구, 친척 혹은 이웃들과 대화를 덜 하게된다
- F12 얼굴을 마주보고 대화할 때도 상대방이 다시 말해 달라고 한다
- P13 목소리가 갈라지고 탁하다
- P14 목소리를 내려면 힘을 주어야 나오는 것 같다
- E15 다른 사람들은 내 음성 문제를 잘 이해하지 못한다고 생각 한다
- F16 음성문제로 개인 생활과 사회생활에 제한을 받는다

- P17 목소리가 언제쯤 맑게 잘 나올지 알 수가 없다(예측이 어렵다)
- P18 목소리를 잘 나오게 하려고 음성을 달리 내보기도 한다
- F19 내 목소리 때문에 대화에 끼지 못하여 소외감을 느낀다
- P20 말할 때는 애를 많이 쓰게 된다
- P21 저녁이 되면 목소리가 더 잠긴다
- F22 음성 문제로 인해 소득(수입)에 감소가 생긴다
- E23 내 목소리 문제로 속이 상한다
- E24 내 목소리 문제로 적극적이지 못할 때가 있다
- E25 음성 문제가 장애로(핸디캡으로) 여겨진다
- P26 말하다가 목소리가 나오지 않아 말을 이을 수 없을 때도 있다
- E27 사람들이 나에게 다시 말해 달라고 할 때 기분이 언짢다
- E28 사람들이 나에게 다시 말해 달라고 할 때 창피함을 느낀다
- E29 목소리 때문에 무능력하게 느껴져 자신감이 떨어진다
- E30 목소리 때문에 수치심을 느낀다

부록 3. ‘노래 부르기’ 문항

- S31 [음성문제 발생 후]⁴⁾ 노래 부르기가 힘들다 0 1 2 3 4
- S32 [음성문제 발생 후] 사람들과 어울릴 때 노래를 잘 부르지 못해 불편할 때가 있다
- S33 [음성문제 발생 후] 누가 나에게 노래를 시킬까봐 걱정을 할 때가 있다

4) 설문 제공 시에는 [괄호] 안의 내용이 없는 문장이 제공되었음. 음성문제와 무관하게 평소 노래 부르기에 문제를 보이는 경우도 있어서, 이를 구별하기 위해 괄호 안 내용을 첨부 수정하였음.

부록 4. SF(Short Form)-36

1) 나는 대체로 건강이 _____

- 1 아주 좋다 2 꽤 좋다 3 좋다 4 괜찮다 5 나쁘다

2) 작년(또는 1년 전)과 비교하여 현재 건강상태는 _____

- 1 아주 좋아졌다
2 다소 좋아졌다
3 같다
4 다소 나빠졌다
5 매우 나빠졌다

* 일상생활 활동 중에 제약을 느끼는 정도를 표시 하십시오. (3번~12번)

- 1= 제한을 많이 느낀다
2= 제한을 약간 느낀다
3= 제한을 전혀 느끼지 않는다

- 3) 달리기 또는 과격한 스포츠를 할 때와 같은 힘찬 운동
4) 책상을 옮기거나 청소기를 밀거나 골프를 치는 것과 같은 중간 정도의 운동
5) 식료품 가게에서 물건을 들거나 옮기는 정도
6) 계단을 몇 개씩 뛰어 오르는 것
7) 계단을 한 개씩 오르는 것
8) 허리 굽히기, 무릎 굽히기
9) 동네 한바퀴를 걷는 것(1500 미터 정도)
10) 학교 운동장을 2-3바퀴 정도 걷는 것(500 미터 정도)
11) 이웃집을 걸어서 가는 것(100미터 정도)
12) 혼자서 목욕을 하거나 옷을 입기

* 지난 한달 동안 신체적인 건강으로 인한 일상생활의 제약이 있었습니까?

(13~16번)

1= 제약이 있었다

2= 제약이 없었다

13) 일 또는 다른 활동을 하는 시간이 줄었다

14) 원하는 만큼 일을 하지 못했다

15) 일을 하거나 다른 활동을 하는 데에 제한을 받았다

16) 일 또는 다른 활동을 어렵게 수행한다(즉, 예전에 비해 힘이 더 든다)

* 지난 한달 동안 정신적인 건강(예, 우울한 느낌 등)으로 인한 일상생활의 제약이 있었습니까? (17~19번)

1= 제약이 있었다

2= 제약이 없었다

17) 일 또는 다른 활동을 하는 시간이 줄었다

18) 원하는 만큼 일을 하지 못했다

19) 평소만큼 조심성 있게 일이나 활동을 하지 못했다

20) 지난 한달 동안 사회활동(예를 들어, 친구 또는 친척의 방문)을 방해하는 신체적 또는 정신적인 문제가 있었습니까?

1 전혀 없었다

2 약간 있었다

3 제법 있었다

4 꽤 많았다

5 아주 많았다

21) 지난 한달 동안 육체적인 고통은 얼마나 있었습니까?

1 없었다

2 아주 약간 있었다

3 약간 있었다

- 4 제법 있었다
- 5 많이 있었다
- 6 아주 많이 있었다

22) 고통으로 인해 일상생활에 방해를 받았습니까?

- 1 전혀 없었다
- 2 조금 있었다
- 3 제법 있었다
- 4 많이 있었다
- 5 아주 많이 있었다

* 지난 한달 동안 기분이 어떤 상태였는지에 대한 질문입니다. (23~31번)

- 1= 항상 느꼈다
- 2= 거의 대부분 느꼈다
- 3= 많이 느낀 편이었다
- 4= 조금 느꼈다
- 5= 별로 느끼지 않았다
- 6= 전혀 느끼지 않았다

23) 활력이 넘쳤습니까?

24) 긴장하여 신경이 쓰이는 일이 있었습니까?

25) 의욕이 전혀 나지 않는 등 감정이 침체되어 있었습니까?

26) 평화롭고 고요한 기분이었습니까?

27) 원기(에너지)가 넘쳤습니까?

28) 낙담하거나 우울하였습니까?

29) 지친 적이 있었습니까?

30) 행복하였습니까?

31) 피곤을 느꼈습니까?

32) 신체적인 또는 정신적인 문제가 사회활동(예를 들어, 친구나 친척을 방문하는 등)을 방해한 시간이 많았습니까?

- 1 항상 그랬다
- 2 대부분 그랬다
- 3 종종 그랬다
- 4 약간 그랬다
- 5 전혀 그렇지 않았다

* 아래의 각 설명에 대해 본인은 어느 정도에 해당되는지 표시하십시오. (33번~36번)

- 1= 정확히 그렇다
- 2= 대부분 그렇다
- 3= 잘 모르겠다
- 4= 대부분 아니다
- 5= 전혀 아니다

- 33) 나는 다른 사람들보다 더 잘 아프다
- 34) 나는 다른 사람보다 건강하다
- 35) 내 건강이 점점 더 나빠지는 것 같다
- 36) 나는 아주 건강하다

부록 5. VHI(Voice Handicap Index)-10

다음의 질문에 대해 본인이 느끼는 증상이 어느 정도인지 숫자에 동그라미(또는 V표)로 표시 하십시오.

0=전혀 그렇지 않다

1=거의 그렇지 않다

2=가끔 그렇다

3=자주 그렇다

4=항상 그렇다

- F1 목소리 때문에 상대방이 내 말을 알아듣기 힘들어한다 0 1 2 3 4
- F3 시끄러운 곳에서는 사람들이 내 말을 이해하기 어려워한다
- P10 사람들이 나에게 목소리가 왜 그러냐고 묻는다
- P14 목소리를 내려면 힘을 주어야 나오는 것 같다
- F16 음성문제로 개인 생활과 사회생활에 제한을 받는다
- P17 목소리가 언제쯤 맑게 잘 나올지 알 수가 없다(예측이 어렵다)
- F19 내 목소리 때문에 대화에 끼지 못하여 소외감을 느낀다
- F22 음성 문제로 인해 소득(수입)에 감소가 생긴다
- E23 내 목소리 문제로 속이 상한다
- E25 음성 문제가 장애로(핸디캡으로) 여겨진다

부록 6. 새로 선택한 VHI 10문항(nVHI-10)

P13 목소리가 갈라지고 탁하다

E23 내 목소리 문제로 속이 상한다

P10 사람들이 나에게 목소리가 왜 그러냐고 묻는다

P14 목소리를 내려면 힘을 주어야 나오는 것 같다

P20 말할 때는 애를 많이 쓰게 된다

E25 음성 문제가 장애로(핸디캡으로) 여겨진다

E24 내 목소리 문제로 적극적이지 못할 때가 있다

P26 말하다가 목소리가 나오지 않아 말을 이을 수 없을 때도 있다

E7 목소리 때문에 타인과 대화를 할 때 긴장을 한다

F16 음성문제로 개인 생활과 사회생활에 제한을 받는다

부록 7. ‘노래 부르기’가 포함된 VHI 10문항(sVHI-10)

S31 [음성문제 발생 후]노래 부르기가 힘들다

P13 목소리가 갈라지고 탁하다

E23 내 목소리 문제로 속이 상한다

P17 목소리가 언제쯤 맑게 잘 나올지 알 수가 없다(예측이 어렵다)

P10 사람들이 나에게 목소리가 왜 그러냐고 묻는다

P14 목소리를 내려면 힘을 주어야 나오는 것 같다

S32 [음성문제 발생 후]사람들과 어울릴 때 노래를 잘 부르지 못해 불편할 때가 있다

P20 말할 때는 애를 많이 쓰게 된다

E25 음성 문제가 장애로(핸디캡으로) 여겨진다

E24 내 목소리 문제로 적극적이지 못할 때가 있다

Abstract

*Korean-Voice Handicap Index(K-VHI): validation of
an original version and development of new version*

Young-Sun Yun

Graduate Program in Speech and Language Pathology, Yonsei University

(Directed by Professor HyangHee Kim)

In evaluations of the severity of voice problems and in order for treatment for dysphonic patients to be effective, clinicians should conduct not only examiner's subjective and objective evaluations, but also a subjective evaluation with the patient. Voice-related psychometric evaluations have been used in recent years for patient's subjective evaluations, but the reliability and validity of this technique, when conducted in the Korean language, have yet to be evaluated.

The VHI(Voice Handicap Index) is one of the most psychometrically robust and well-studied instruments among the various instruments for measuring quality of life in patients suffering from voice disorders. However, the VHI has yet to be established as a gold standard for such testing. The primary objectives of this study were 1) to validate the Korean-translated versions of the VHI-30 and the VHI-10, and 2) to develop a new VHI version considering the cross-cultural differences. One hundred fifty-six Korean adult patients, all of whom had symptoms of dysphonia without laryngeal cancer, as well as 66

normal Korean adults participated in this study. The survey was conducted using the VHI-30, three 'singing' items, the IPVI(Iowa Patient's Voice Index)-(1) and the SF(Short Form)-36. Acoustic analysis with the MDVP (Multi-Dimensional Voice Program) was conducted only in the patient group. The results were as follows.

1. With regard to the reliability of the VHI-30, Pearson correlations in test-retest were very high($r=0.971$, $p<0.01$). Cronbach's α in internal consistency was 0.969. The control group received significantly lower scores than did the patient group. For the concurrent validity of the VHI-30, with regard to the total score of the patient group, Spearman correlations between the VHI-30 and the IPVI-(1) were high($r=0.664$, $p<0.01$), and the correlations between the VHI-30 and the SF-36 were low($r=0.372$, $p<0.01$). With regard to the construct validity, the VHI-30 was divided into three components via factor analysis.

2. Patients with neurogenic voice disorders evidenced worse VHI-30 scores than those suffering from any other voice disorders(functional, structural, and inflammatory disorders) and the controls. Those with neurogenic laryngeal disease evidenced the most severe impairments, supporting the conclusions of previous studies. This result suggests that a group study of laryngeal diseases could be achieved using the VHI-30.

3. Correlations between the VHI-30 scores and the results of MDVP parameters were in relatively weak agreement. The VHI-30 and acoustic measurements provided independent information supporting the results of previous studies.

4. Also for the VHI-10, statistical analyses revealed high reliability and validity values. The correlations of total scores between the VHI-30 and the VHI-10 were high($r=0.968$, $p<0.01$). The mean ratio of total scores between the VHI-10 and the VHI-30 were 0.368. The items on the VHI-10 were more

sensitive than those of the VHI-30, supporting the results of previous studies. VHI-10 had a single component according to the results of factor analysis.

5. Ten questionnaire items as a 'new VHI-10'(nVHI-10) items were selected from the VHI-30. Five items from the original VHI-10 were remained in the nVHI-10. The mean ratio of the nVHI-10 to the VHI-30 was 0.408. The items on the nVHI-10 were more sensitive than those of the VHI-10.

6. In order to reflect the cross-cultural differences, 'singing VHI-10' (sVHI-10) was developed by including eight items of the VHI-30 and 2 singing items (S31 as a physical factor: 'My voice problem makes it difficult for me to sing', and S32 as a functional factor: 'My singing difficulties restrict social life'). The mean ratio of the sVHI-10 to the VHI-30 was 0.456. The sVHI-10 were more sensitive than those of the VHI-10 and the nVHI-10. These results confirmed that cross-cultural differences exist with regard to the selection of survey items.

The validity and reliability of VHI-30 and VHI-10 questionnaires are maintained even after translating into Korean. In order to conduct the voice-related psychometric evaluation questionnaire or to decipher its results, cross-linguistic and cross-cultural differences should be considered.

Key words: voice disorders, VHI(Voice Handicap Index), quality of life, SF-36, cross-cultural research