파킨슨증으로 인한 마비말장애의 청지각적 평가에 대한 신뢰도 연구

A Reliability Study on the Auditory-perceptual Evaluation of Parkinsonian Dysarthria

김 향 희* **· 이 미 숙*· 김 선 우*· 이 원 용***
Hyanghee Kim · Mi sook Lee · Sun woo Kim · Won yong Lee

ABSTRACT

An auditory-perceptual evaluation has long been utilized in assessing dysarthric speech. The process involves subjective judgement and the results might vary depending on clinical experiences or training of listeners. This study aimed to investigate reliability of the auditory-perceptual evaluation of 22 multi-dimensional variables on 6 patients with Parkinsonian speech disorders. Listeners were divided into two groups: one consisted of 6 speech therapists with clinical experiences for three years or more, and the other 6 graduate students without any previous clinical background. The results showed that the former evaluated dysarthric speech with higher inter-rater and intra-rater reliabilities than the latter. Furthermore, such speech variables as 'precise consonant,' 'speech intelligibility,' and 'SMR regularity' were more influenced than others by clinical experiences. We, therefore, postulated that a reliable auditory-perceptual evaluation of dysarthric speech may require adequate amount of clinical training of listeners.

Keywords: Auditory-perceptual evaluation, Parkinsonian dysarthria, Reliability, Clinical experience

1. 서 론

말장애에 대한 청지각적 평가란 청자가 말장애 환자의 말을 귀로 듣고 평가척도(rating scale 또는 scaling procedure)상에 명기된 기준에 따라 환자의 말 특징, 장애의 중증도 등을 판정하는 평가방법이다. 마비말장애(dysarthria)와 같은 말운동장애(motor speech disorder)의 경우, 일반적인 운동조절기능이나 구강기제구조 등을 관찰하는 것만으로는 환자의 말 산출 기능을 정확히 판단할 수 없기 때문에 자발화시 나타나는 말의 유형 및 정도를 평가하는데 있어서 청지각적 판단에 많이 의존하고 있다(Darley, Aronson & Brown, 1969a, b; Logemann, Fisher, Boshes & Blonsky, 1978). 더욱이, Moll (1964)은 모든 말 문제를 평가할 때 청자의 청지각적 판단이 가장 바람직하고 필수적인 방법이라고 제안하였고, Moran & Gilbert (1984)는 음성평가의 경우에도 음향학적 또는 공기역

* 연세대학교 대학원 언어병리학 협동과정
** 연세대학교 의과대학 재활의학교실
*** 성균관대학교 의과대학 삼성서울병원 신경과학교실

학적 측정에 비해 청지각적 평가가 더 유용하다고 주장한 바 있다.

청지각적 평가 시 사용하는 '평가척도'는 청자가 들은 것을 판정하기 위하여 일정한 기준을 제시하는데, 이는 말장애를 포함한 의사소통 장애의 다양한 변수(parameter)들을 직접적으로 평가하는데 적용되어 왔다. 청지각적 평가는 환자의 말 특징을 질적으로 평가할 수 있는 반면, 청자나 평가 환경 등의 영향을 많이 받기 때문에 기기를 사용하는 음향학적 방법에 비하여 객관성이 결여될 수 있다. 특히, 평가자의 주관적인 판단에 많이 의존하므로 임상경험이나 훈련정도에 따라 평가자내(intra-rater) 또는 평가자간(inter-rater) 신뢰도가 영향을 받을 수 있다. 따라서 청지각적 평가의 신뢰도를 높이기 위해서는 개개의 말 특징자체에 대한 지각적 이해, 동시에 나타나는 여러 요소들로부터 하나의 요소를 구별하는 지각적 능력, 말이나 음성 자료를 평가하기 위한 기술 및 척도의 동일성, 판정에 대한 일관적인 신뢰도, 평가자간 차이의 최소화 등이 선행되어야 한다고 주장되었다(Kent, 1996). 더욱이, 청자에게 별도의 청지각적 훈련을 실시할 경우에는 평가자간 신뢰도가 더 높아지는 경향이 있다고 보고 된 바 있다(McWilliams, 1954; McWilliams & Philips, 1979). 특히, Chan & Yiu (2002)의 연구에서는 청자의 훈련정도에 따라 음성 평가의 신뢰도가 달라지며, 특히 거친 소리(roughness)와 기식음(breathiness)의 경우 훈련에 따른 신뢰도 향상이 두드러지게 나타났다.

마비말장애의 전체적인 말 특징을 대상으로 한 신뢰도 연구는 그다지 많지 않다. 또한 몇몇의 연구조차도 청자의 임상경험에 따른 청지각적 판단의 차이를 규명하기보다는 한 종류의 평가자군에서 말척도에 대한 평가자간 및 평가자 내 신뢰도를 살펴본 것이 대부분이다. Ozsancak 등 (2002)은 9명의 치료사를 대상으로 10 명의 마비말장애 환자에 대한 청지각적 평가를 실시함으로써 척도의 신뢰성을 검증하였는데, 총 33 개 항목을 5점 척도로 평가한 결과 평가자내 신뢰도는 90%, 평가자간 신뢰도는 80%로 밝혀졌다. Whitehill 등 (2003)은 언어병리학을 전공하였으나 마비말장애 분야를 담당하지 않는 10 명의 치료사를 대상으로 파킨슨병 환자 19 명에 대한 청지각적 판단력을 알아보았는데, 7 점 척도로 평가한 21 개 항목에 대해 평가자내 신뢰도는 94.52%, 평가자간 신뢰도는 88%로 비교적 높게 나타났다.

본 연구에서는 파킨슨증(Parkinsonism) 환자들을 대상으로 평가척도에 따른 청지각적 평가를 실시하면서 평가자의 임상경험에 따른 신뢰도의 차이를 알아보고, 각 말 특징 변인들을 평가할 때 관찰되는 신뢰도를 측정하고자 한다.

2. 연구 방법

2.1 연구 대상

치료경험 여부에 따른 신뢰도의 차이를 알아보기 위해 평가자를 크게 두 집단으로 분류하였다. 첫 번째 집단은 언어치료사 자격증을 소지하고 마비말장애 환자에 대한 평가 및 치료 경험이 3 년 이상(범위=3 년~7 년 4 개월; 평균=3 년 11 개월)인 6 명의 언어치료사를 포함하였다. 두 번째 집단은 언어병리학 대학원 과정에 재학 중이나, 마비말장애 수업을 수강한 적이 없는 6 명의 학생들로 구성되었다. 이들은 각각 다계통위축증(multiple system atrophy, MSA) (총 3 명, 남 : 녀=1 : 2;

평균연령=67 세)과 특발성 파킨슨병(idiopathic Parkinson's disease, IPD) (총 3 명, 남:녀=2 : 1; 평균연령=63.3 세) 환자[1] 6 명의 말 자료를 듣고 평가척도[2]에 준하여 각 변인들을 평가하였다. 환자군은 모두 신경과에서 확진된 파킨슨증 환자들이며, Hoehn & Yahr(H & Y) stage가 II보다 심하고 MMSE 점수에서 20 점 이상을 획득한 환자들이었다.

2.2 연구 방법

2.2.1 자료수집

각 환자에게 대화과제, 그림설명과제, 모음연장과제, DDK(DiaDochoKinetic) 속도과제(즉, AMR 및 SMR 과제) 등을 실시하여 카세트 녹음기(Marantz PMD221)에 각 반응을 녹음하였다. 녹음은 소음이 적은 조용한 치료실에서 진행되었다.

2.2.2 평가항목

총 22 개 항목을 대상으로 평가를 실시하였는데, 음성 단계에서는 '발성부전(hypophonia),' '음도수준(pitch level),' '단조로움(monotone),' '쉰소리(hoarseness),' '기식음(breathiness),' '거친 소리(harshness),' '음성진전(voice tremor)' 등 총 7개 항목이 포함되었다. 호흡단계에서는 /아/ 모음의 최대발성지속시간(maximum phonation time, MPT)을, 공명단계에서는 비음도(nasality)을 평가하였다. 말속도 측면의 경우, '교대운동속도(alternate motion rate, AMR)과제'는 /퍼/, /터/, /커/의 세 음절을, '일련운동속도(sequential motion rate, SMR)과제'는 /퍼터커/ 음절을 5 초간 반복한 속도 및 규칙성을 판단하였다. 이밖에, '점차로 느려지는 발화(bradyphemia)[3],' '점차로 빨라지는 발화(tachyphemia)[4],' '부적절한 쉼(inappropriate silences)'을 함께 평가하였다. 조음 측면에서는 '자음정확도(precise consonant)'와 '말명료도(speech intelligibility)' 항목을 포함하였다.

2.2.3 평가절차

실험에 참가한 총 12 명의 측정자는 모든 말 자료를 듣고 평가척도에 따라 각 변인들을 평가하였으며, 평가자 간 합의(consensus rating)는 허용되지 않았고, 별도의 청지각훈련 역시 실시하지 않았다. 평가자 내 신뢰도를 알아보기 위해서는 동일한 말자료에 대해 2 주 간격으로 초기평가와 재평가를 시행하였다.

각 항목별로 세부적인 절차를 살펴보면, 'hypophonia'는 대화 및 그림설명 과제를 듣고 판단하였다. 'Pitch level'은 모음연장 및 대화 과제에서 음도의 높낮이를 평가하였고, 'monotone'은 대화 및 그림설명 과제에서 평가하였다. 'Hoarseness'는 청지각적으로 쉰 듯한 인상을 주는 정도를, 'breathiness'

---

1) 본 연구자들에 의한 선행연구논문 [김향희 등 (2003, 2004)]의 대상군 중에서 선택되었음.

2) 본 연구자들에 의한 선행연구논문 [김향희 등 (2004)]의 평가척도에 'pitch level'과 'inappropriate silence'를 추가하였으며, 'hypernasality'와 'hyponasality'를 합하여 'nasality'로 명명하였음.

3) 본 연구자들에 의해 제안된 용어임;
'brady-'는 '느린(slow)-'의 의미를 지님 (ref. The American Heritage Dictionary, 4th ed., 2000, Boston: Houghton Mifflin Co.).

4) 'tachy-'는 '점차로 빨라지는(accelerated)-'의 의미를 지님 (ref. The American Heritage Dictionary, 4th ed., 2000, Boston: Houghton Mifflin Co.).

는 발화 시 공기가 새어나오는 듯한 인상을 주는 정도를, 그리고 'harshness'는 불규칙한 진동(jitter와 shimmer)을 가진 듯한 인상을 주는 정도를 평가하였다. 'Voice tremor'는 모음연장 및 자발화, 읽기 과제를 대상으로 하였으며, 'MPT'는 초시계를 이용해 모음 /아/를 최대한 길게 발성한 시간을 측정하되, 과제를 2 회 이상 시행한 경우에는 가장 길게 발성한 시간을 선택하였다. 'Nasality'는 대화 및 읽기, 모음연장 과제에서 '0'(정상)을 기준으로 '−'는 과소비성(hyponasality), '+'는 과대비성(hypernasality)의 정도를 청지각적으로 평가하였다.

'AMR' 항목은 /퍼/, /터/, /커/ 음절을, 'SMR'은 /퍼터커/ 음절을 각각 5초 동안 반복한 횟수 및 규칙성 정도를 평가하였는데, 과제수행 시간이 5 초 미만인 경우에는 실시하지 않은 것으로 간주하였다. 말의 뭉침(short rushes)이 심한 경우에는 반복 횟수를 계산하지 않았으며, 규칙성 역시 평가할 수 없는 것으로 간주하였다. 반복이 경미하게 뭉쳐서 나타날 경우에는 환자의 반복 속도에 비례하여 횟수를 계산하였으며, 매우 심하게 소리의 뭉침이 나타날 경우에는 뭉침 부분을 제외하고 반복 횟수를 산정하였다. 거의 대부분의 음절이 뭉친 경우에는 반복 횟수의 계산은 제외하고 규칙성만을 평가하였다. 대화과제에서 측정한 'bradyphemia'와 'tachyphemia' 항목들에서는 환자의 말속도가 점차로 느려지는 현상을 보이는 경우에는 'tachyphemia'항목에서 '0'(정상)으로 평가하고, 점차 빨라지는 경우에는 'bradyphemia' 항목을 '0'(정상)으로 평가하였다. 'Inappropriate silences'는 대화과제에서 나타나는 부적절한 쉼을 측정하였는데, 이어지는 음절이나 어절, 구, 문장 등을 산출하기 전에 '음' 등의 삽입어를 첨가하는 경우, 그리고 내용을 이해하지 못해 시간적 단절이 발생하는 경우는 제외하였다. 'Precise consonant' 및 'speech intelligibility'는 발화 전사를 통해 비율을 산정하지 않고 자음의 정확한 정도(자음정확도)와 청자가 알아들은 정도(명료도)를 청지각적으로 평가하였다.

2.2.4 자료분석

평가자간 신뢰도는 평가척도 내 각 항목별로 평가자들 중 가장 많은 평가자가 선택한 값(예: 0, 1, 2, 3, 4)의 개수를 기준으로 서열화하였다. 평가자내 신뢰도는 초기평가와 재평가 간 값이 일치한 개수를 기준으로 산정한 후 이들을 각각 높은 순으로 서열화하였다. 통계처리는 Windows용 SPSS (v.12.0) 프로그램을 사용하였으며, 스피어만의 순위 상관관계(Spearman's rank-order correlation)를 통해 두 집단 간 신뢰도 순위의 상관성을 알아보았다.

3. 결과 및 고찰

3.1 평가자간 신뢰도

표 1에서 볼 수 있듯이, 각 집단 내에서 평가자간 평균신뢰도는 치료사 집단이 72.47% (범위: 97.22~52.78%), 학생 집단이 69.70% (범위: 86.11~63.89%)로 치료사 집단의 평가가 약간 더 높은 수치를 보였다. 항목별 신뢰도수치를 살펴보면, 두 집단에서 공통적으로 높은 신뢰도를 보인 항목은 MPT (치료사 97.22%; 학생 86.11%), tachyphemia (88.89%; 88.89%), AMR /퍼/ 반복횟수 (83.33%; 88.89%), AMR /터/ 반복횟수 (83.33%; 83.33%), AMR /커/ 반복횟수 (83.33%; 86.11%)

등을 포함한다. MPT나 AMR /퍼/, /터/, /커/ 반복횟수의 경우, 초시계로 시간을 측정하여 평가자들이 정해진 시간 내에 반복된 음절수를 세는 비교적 객관적인 방법을 사용했기 때문에 평가자들 간 일치도가 높은 것으로 해석할 수 있다. 각 집단 내의 평가자간 신뢰도 순위에서 상위에 위치한 변인일수록 일반적으로 평가에 용이한 항목이라고 간주할 수 있다. 반면에, 하위에 위치한 변인들은 청지각적인 판단에 어려움을 겪는 경우에 해당된다. 그런데, 신뢰도 수치상에서 두 집단이 서로 다른 양상을 보이는데 치료사 집단의 경우 70% 이상의 신뢰도를 보이는 항목은 상위 13 개 항목인데 비하여, 학생 집단의 경우는 상위 7 개로 훨씬 적은 수였다.

그런데, 두 집단 모두에서 'breathiness,' 'harshness,' 'pitch level,' 'hoarseness,' 'voice tremor,' 'monotone' 등과 같이 음성의 질과 관련된 평가항목들은 일반적으로 낮은 신뢰도를 보였는데, 이는 아마도 음성의 질에 대한 항목들은 다른 항목들에 비하여 평가자의 청지각적 판단력에 더 많이 의존하는 요인이기 때문인데(Kreiman et al., 1992), 음성치료 등의 전문화된 영역에서 활동하는 언어치료사들이나 이비인후과 의사들조차도 일관적이지 못한 측정결과를 보인다고 알려져 있다(Yumoto et al., 1984). 이처럼, 치료사 및 학생집단이 특정 평가항목에 대하여 공통적으로 수월하게 느끼거나 혹은 공통적으로 어려움을 겪는 항목들은 평가자의 치료경험에 따른 영향을 비교적 적게 받는다고 결론지을 수 있다.

한편, 'precise consonant,' 'speech intelligibility,' 'SMR 규칙성' 등은 학생집단에 비해 치료사집단에서 더 높은 신뢰도를 보인 항목들 (신뢰도 차이: 15%<)로서, 이들을 평가할 때 평가자의 치료경험 여부가 영향을 미친다는 것을 알 수 있다. 치료사집단과 학생집단 간에 가장 큰 신뢰도 차이(30.56%)를 보였던 'precise consonant'는 두 집단의 신뢰도 순위가 각각 3 위와 22 위로 순위 차이 역시 가장 크게 나타났다. 'Speech intelligibility'는 치료사집단과 학생집단의 신뢰도 차이가 22.22%였으며, 이들의 순위는 4 위(치료사)/15 위(학생)로 비교적 큰 차이를 보였다. 'Speech intelligibility'를 결정하기 위해서는 말속도, 운율, 그리고 자음정확도 등의 다중변인들(Weismer & Martin, 1992)에 대한 판단이 동시에 이루어져야 하므로 임상경험이 큰 영향을 미칠 수 있다. 'Precise consonant'도 각각의 자음에 대한 명확한 조음여부를 판단하는 항목으로서 풍부한 임상경험을 필요로 한다. 'SMR 규칙성'의 경우, 연속음절인 /퍼터커/이 얼마나 규칙적인지를 평가해야 하므로 'AMR 규칙성'에 비해 학생 집단에서 판단에 더 어려움을 겪었던 것으로 보인다. 특히, 파킨슨증을 보이는 환자들이 조음속도가 빠르고 음절 간 뭉침현상이 관찰되므로 (김향희 등, 2003) 규칙성의 판단이 쉽지는 않으리라 생각된다.

표 1. 집단별 각 항목에서의 평가자간 신뢰도

항 목	순위(신뢰도, %)		순위 차이(A-B) (신뢰도 차이, %)
	치료사 집단(A)	학생 집단(B)	
MPT	1 (97.22)	3 (86.11)	-2 (11.11)
tachyphemia	2 (88.89)	1 (88.89)	1 (0.00)
precise consonant	3 (86.11)	22 (55.56)	-19 (30.56)*
AMR /퍼/ 반복횟수	4 (83.33)	1 (88.89)	3 (-5.56)
AMR /터/ 반복횟수	4 (83.33)	5 (83.33)	-1 (0.00)
AMR /커/ 반복횟수	4 (83.33)	3 (86.11)	1 (-2.78)
inappropriate silence	4 (83.33)	8 (69.44)	-4 (13.89)
speech intelligibility	4 (83.33)	15 (61.11)	-11 (22.22)*
SMR 반복횟수	9 (75.00)	6 (80.56)	3 (-5.56)
SMR 규칙성	9 (75.00)	20 (58.33)	-11 (16.67)*
nasality	11 (72.22)	6 (80.56)	5 (-8.33)
AMR /퍼/ 규칙성	11 (72.22)	15 (61.11)	-4 (11.11)
bradyphemia	11 (72.22)	10 (66.67)	1 (5.56)
breathiness	14 (69.44)	8 (69.44)	6 (0.00)
AMR /터/ 규칙성	14 (69.44)	15 (61.11)	-1 (8.33)
hypophonia	16 (66.67)	15 (61.11)	1 (5.56)
harshness	17 (58.33)	15 (61.11)	2 (-2.78)
pitch level	18 (55.56)	11 (63.89)	7 (-8.33)
hoarseness	18 (55.56)	20 (58.33)	-2 (-2.78)
voice tremor	18 (55.56)	11 (63.89)	7 (-8.33)
AMR /커/ 규칙성	18 (55.56)	11 (63.89)	7 (-8.33)
monotone	22 (52.78)	11 (63.89)	11 (-11.11)
평균 신뢰도	72.47	69.70	2.78

* : 신뢰도 차이(A 신뢰도-B 신뢰도)가 15% 이상인 경우

두 집단 간의 평가자간 신뢰도 순위의 상관성을 분석한 결과, 상관관계 계수 ρ는 0.428 (p=.047)로 치료사 집단과 학생 집단 간에 다소 낮은 상관관계를 보이고 있다. 그림 1에서 보듯이, 두 집단의 분포도는 대각선을 중심으로 넓게 퍼져있어 집단별로 수월하게 평가할 수 있는 변인들이 다르다는 것을 의미하며, 따라서 임상경험 여부가 평가자 간 신뢰도에 영향을 미치는 요소임을 시사하고 있다.

3.2 평가자 내 신뢰도

평가자 내 신뢰도를 분석한 결과, 치료사집단의 평균 신뢰도수치는 67.93%, 학생집단은 63.76%로 치료사집단의 평가자내 신뢰도가 다소 높게 나타났다(표 2 참조). 치료사집단의 경우는 신뢰도 범위가 94.44~47.22%였으며, 학생집단의 경우는 94.44~52.78%로서 두 집단 모두에서 앞서 제시한 평가자간 신뢰도보다 그 범위가 훨씬 넓었다. 항목별 신뢰도수치를 살펴보면, 두 집단에서 공통적으로 높은 신뢰도를 보인 항목은 MPT (치료사 94.44%; 학생 94.44%), tachyphemia (86.11%;

91.67%), AMR /퍼/ 반복횟수 (75.00%; 83.33%) 등을 포함한다. MPT나 AMR /퍼/ 반복횟수의 경우, 초시계로 시간을 측정하여 정해진 시간 내에 반복된 음절수를 세는 비교적 객관적인 방법을 사용했기 때문에 평가자들이 재평가를 하더라도 일치도가 높을 수 있었다.

그리고, 'hypophonia,' 'breathiness,' 'harshness,' 'hoarseness,' 그리고 'voice tremor' 등과 같이 음성의 질과 관련된 평가항목들에서는 두 집단 모두가 공통적으로 60%대 이하의 낮은 평가자내 신뢰도를 보였다. 한편, 평가자내 신뢰도가 학생집단과 치료사집단간에 15% 이상의 신뢰도 차이를 보였던 항목들은 harshness (27.78% 차이)와 AMR /퍼/ 규칙성 및 AMR /터/ 반복횟수 (각각 16.67% 차이)였다. 이 중에서 AMR /터/ 반복횟수는 앞서 제시된 평가자간 신뢰도의 경우 두 집단의 신뢰도수치가 모두 83.33%로 그 차이가 '0'였음에도 불구하고, 평가자내 신뢰도에서 집단간 이처럼 비교적 큰 차이를 보인 것은 흥미로운 결과이나 그에 대한 해석은 불분명하다.

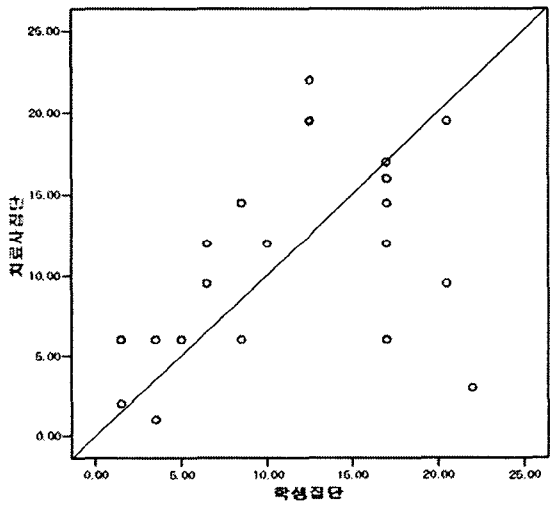

그림 1. 두 집단 간 평가자간 신뢰도 순위의 상관관계

두 집단 간 평가자 내 신뢰도 순위의 상관관계 계수 ρ는 0.72로 비교적 높은 수준의 상관성을 보이는 것으로 분석되었다(p=.000). 그림 2를 통해 두 집단 간 평가자내 신뢰도가 대각선 축을 중심으로 직선에 가깝게 위치하고 있다는 것을 확인할 수 있다. 이는 평가자가 동일한 항목에 대해 재평가를 실시할 때 집단 또는 치료경험에 관계없이 모든 평가자가 유사한 정도의 수월함 또는 어려움을 보인다는 것을 의미한다.

표 2. 각 항목에서의 집단별 평가자내 신뢰도

항목	순위(신뢰도,%)		순위 차이(A-B)
	치료사 집단(A)	학생 집단(B)	(신뢰도 차이, %)
MPT	1 (94.44)	1 (94.44)	0 (0.00)
AMR /터/ 반복횟수	2 (86.11)	7 (69.44)	-5 (16.67)*
tachyphemia	2 (86.11)	2 (91.67)	0 (-5.56)
nasality	4 (80.56)	7 (69.44)	-3 (11.11)
pitch level	5 (77.78)	5 (72.22)	0 (5.56)
AMR /퍼/ 반복횟수	6 (75.00)	3 (83.33)	3 (-8.33)
AMR /커/ 반복횟수	7 (72.22)	10 (66.67)	-3 (5.56)
bradyphemia	7 (72.22)	12 (63.89)	-5 (8.33)
SMR 반복횟수	9 (69.44)	4 (77.78)	5 (-8.33)
precise consonant	9 (69.44)	13 (55.56)	-4 (13.89)
inappropriate silence	11 (66.67)	5 (72.22)	6 (-5.56)
monotone	12 (63.89)	10 (66.67)	2 (-2.78)
AMR /퍼/ 규칙성	12 (63.89)	21 (47.22)	-9 (16.67)*
AMR /터/ 규칙성	12 (63.89)	16 (52.78)	-4 (11.11)
SMR 규칙성	12 (63.89)	13 (55.56)	-1 (8.33)
speech intelligibility	12 (63.89)	19 (50.00)	-7 (13.89)
hypophonia	17 (58.33)	19 (50.00)	-2 (8.33)
breathiness	17 (58.33)	7 (69.44)	10 (-11.11)
harshness	17 (58.33)	22 (30.56)	-5 (27.78)*
hoarseness	20 (55.56)	16 (52.78)	4 (2.78)
voice tremor	21 (47.22)	13 (58.33)	8 (-11.11)
AMR /커/ 규칙성	21 (47.22)	16 (52.78)	5 (-5.56)
평균 신뢰도	67.93	63.76	4.17

* : 신뢰도 차이(A 신뢰도-B 신뢰도)가 15% 이상인 경우

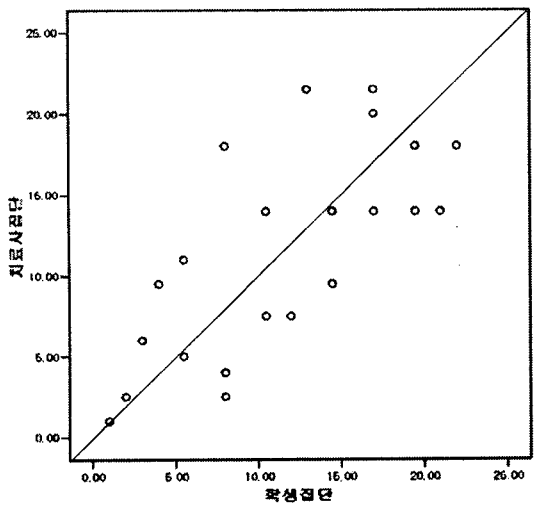

그림 2. 두 집단의 평가자내 신뢰도 순위의 상관관계

4. 결 론

본 연구에서는 다계통위축증과 특발성 파킨슨병 환자들의 마비말장애의 변인들을 청지각적으로 평가할 때 변인들의 유형과 평가자들의 임상경험 등에 따라 평가자간 및 평가자내 신뢰도가 달라지므로 그 요인들을 고려하는 것이 중요하다고 밝히고 있다. 특히 평가자들의 임상경험과는 상관없이 음성의 질에 대한 변인들은 치료사와 학생집단이 공통적으로 낮은 신뢰도수치를 보임으로 특히 많은 훈련을 통한 정확한 평가에 접근해야 할 것이다.

참 고 문 헌

김은정. 2003. 정상 청년층과 노년층의 최대발성시간 및 조음교대운동속도 비교. 연세대학교 대학원 석사학위논문.

김향희, 이미숙, 김선우, 이원용. 2003. "파킨슨병과 다계통위축증 환자군 간의 말속도 비교평가." *음성과학*, 10(4), 221-228.

김향희, 이미숙, 김선우, 최성희, 이원용. 2004. "파킨슨증으로 인한 마비말장애에 대한 청지각적 평가척도." *음성과학*, 11(2), 39-49.

이영미, 김향희. 2001. "대화와 그림설명을 통한 한국 성인발화의 비교분석." *언어청각장애연구*, 6, 40-52.

Darley, FL., AE. Aronson, & JR. Brown, 1969a. "Differential diagnostic patterns of dysarthria." *Journal of Speech & Hearing Research*, 12, 246-269.

Darley, FL., AE. Aronson, & JR. Brown, 1969b. "Clusters of deviant speech dimensions in the dysarthrias." *Journal of Speech & Hearing Research*, 12, 462-469.

Chan, KMK, & MLE. Yiu, 2002. "The effect of anchors and training on the reliability of perceptual voice evaluation." *Journal of Speech, Language, and Hearing Research*, 45, 111-126.

Kent, RD., 1996. "Hearing and believing: some limits to the auditory-perceptual assessment of speech and voice disorders." *American Journal of Speech-Language Pathology*, 5, 7-23.

Kreiman, J., BR. Gerratt, K. Precoda, & GS. Berke, 1992. "Individual differences in voice quality perception." *Journal of Speech & Hearing Research*, 35, 512-520.

Logemann, J., H. Fisher, B. Boshes & E. Blonsky, 1978. "Frequency and co-occurrence of vocal tract dysfunctions in speech of a large sample of Parkinson's patients." *Journal of Speech and Hearing Disorders*, 43, 47-57.

McWilliams, BJ., 1954. "Some factors in the intelligibility of cleft palate speech." *Journal of Speech and Hearing Disorders*, 19, 524-527.

McWilliams, BJ., & BJ. Philips, 1979. "Velopharyngeal incompetence." Audio Seminars in Speech Pathology. Philadelphia: WB Saunders.

Moll, KL., 1964. "Objective measures of nasality." *Cleft Palate Journal*, 1, 371-374.

Moran, ALJ, & HIL. Gilbert, 1984. "Relation between voice profile ratings and aerodynamic and acoustic parameters." *Journal of Communication Disorders*, 17, 245-260.

Ozsancak, C., AM. Parais, & P. Auzou, 2002. "Perceptual analysis of dysarthria: Presentation and validation of a clinical scale. Priminary study (French)." *Revue Neurologique*, 4, 158(4),

431-438.

Whitehill, TL., JK. Ma, & AS, Lee. 2003, "Perceptual Characteristics of Cantonese Hypokinetic Dysarthria." *Clinical Linguistics & Phonetics*, 17(4-5), 265-271.

Weismer, G., & RE. Martin, 1992. "Scaling procedures for the measurement of speech intelligibility." In RD. Kent (ed.), Intelligibility in Speech Disorders (pp. 67-118). Philadelphia: John Benjamins Publishing Co.

Yumoto, E., Y. Sasaki, & H. Okamura, 1984. "Harmonics-to-noise ratio and psychophysical measurement of the degree of hoarseness." *Journal of Speech and Hearing Research*, 27, 2-6.

접수일자: 2004. 11. 01
게재결정: 2004. 11. 30

▲ 김향희
서울특별시 서대문구 신촌동 134번지 (우: 120-752)
연세대학교 대학원 언어병리학 협동과정 교수
연세대학교 의과대학 재활의학교실 교수
Tel: +82-2-361-7537, Fax: +82-2-6748-7578
E-mail: hkim@yumc.yonsei.ac.kr

▲ 이미숙
서울특별시 서대문구 신촌동 134번지 (우: 120-752)
연세대학교 대학원 언어병리학 협동과정
Tel: +82-2-361-7578, Fax: +82-2-6748-7578
E-mail: camusms@hanmail.net

▲ 김선우
서울특별시 서대문구 신촌동 134번지 (우: 120-752)
연세대학교 대학원 언어병리학 협동과정
Tel: +82-2-361-7578, Fax: +82-2-6748-7578
E-mail: weddingdayre@hanmail.net

▲ 이원용
서울특별시 강남구 일원동 50번지 (우: 135-710)
성균관대학교 의과대학 삼성서울병원 신경과학교실 교수
Tel: +82-2-34101-3593, Fax: +82-2-3410-0052
E-mail: ywlee@smc.samsung.co.kr

<부 록>

Speech Rating Scale

1. hypophonia

0 = normal (정상임)
1 = mild (약간 작은 소리)
2 = moderate (매우 작은 소리를 자주 내며, cue를 주면 크게 낼 수 있음)
3 = severe (매우 작은 소리를 항상 내며, cue를 주어도 크게 낼 수 없음)
4 = profound (전혀 들리지 않고 입 모양만 냄)

2. pitch level (대화 수준)

-4 = profoundly low
-3 = severely low
-2 = moderately low
-1 = mildly low
0 = normal
1 = mildly hight
2 = moderately high
3 = severely high
4 = profoundly high

3. monotone

0 = normal
1 = mildly monotonous
2 = moderately monotonous
3 = severely monotonous
4 = profoundly monotonous

4. hoarseness (쉰 듯한 소리)

0 = normal
1 = mild
2 = moderate
3 = severe
4 = profound

5. breathiness (공기가 새는 듯한 소리)

0 = normal
1 = mild
2 = moderate
3 = severe
4 = profound

6. harshness (거친 소리)

0 = normal
1 = mild
2 = moderate
3 = severe
4 = profound

7. voice tremor

0 = normal (정상임)
1 = mild (경미한 정도로 가끔 관찰됨)
2 = moderate (경미한 정도로 자주 관찰됨)
3 = severe (심한 정도로 자주 관찰됨)
4 = profound (심한 정도로 항상 관찰됨)

8. MPT (/아/를 최대한 길게 발성한 시간)[1]

-청년층: 16.5±4.4 (단위: 초)
0 = normal (12.1 이상)
1 = mildly short (7.6-12.0)
2 = moderately short (3.1-7.5)
3 = severely short (3.0 이하)
4 = profound (전혀 소리를 내지 못함)

-노년층: 12.8±3.6 (단위: 초)
0 = normal (9.2 이상)
1 = mildly short (5.5-9.1)
2 = moderately short (1.8-5.4)
3 = severely short (1.7 이하)
4 = profound (전혀 소리를 내지 못함)

9. nasality

-4 = profound hyponasality
-3= severe hyponasality
-2= moderate hyponasality
-1= mild hyponasality
0 = normal
1 = mild hypernasality
2 = moderate hypernasality
3 = severe hypernasality
4 = profound hypernasality

AMR (조음 정확성을 유지하면서 가능한 빨리, 규칙적으로 5 초 동안 /퍼/, /터/ 또는 /커/를 반복)[2,3]

10-11. AMR /퍼/ 반복횟수 및 규칙성

. 반복횟수

-청년층: 33.1±3.2 (단위: 회)
-4 = profoundly slow (6.3 이하)
-3 = severely slow (6.4-13)
-2 = moderately slow (13.1-19.7)
-1 = mildly slow (19.8-26.4)
0 = normal (26.5-39.7)
1 = mildly fast (39.8-43.1)
2 = moderately fast (43.2-46.5)
3 = severely fast (46.6-49.9)
4 = profoundly fast (50 이상)

-노년층: 27.7±3.0 (단위: 회)
-4 = profoundly slow (4.9 이하)
-3 = severely slow (5-10.6)
-2 = moderately slow (10.7-16.3)
-1 = mildly slow (16.4-22)
0 = normal (22.1-33.3)
1 = mildly fast (33.4-36.2)
2 = moderately fast (36.3-39.1)
3 = severely fast (39.2-42)
4 = profoundly fast (42.1 이상)

· 규칙성
0 = normal (정상)
1 = mild(약간 불규칙하나 음절 간 뭉침이 없음)
2 = moderate(자주 불규칙하나 음절 간 뭉침이 없음)
3 = severe (대부분 불규칙하며 음절 간 뭉침이 간헐적으로 관찰됨)
4 = profound (대부분 불규칙하며 음절 간 구분이 안 됨)

12-13. AMR /터/ 반복횟수 및 규칙성

· 반복횟수

-청년층: 34.1±3.9 (단위: 회)
-4 = profoundly slow (7.3 이하)
-3 = severely slow (7.4-14.0)
-2 = moderately slow (14.1-20.7)
-1 = mildly slow (20.8-27.4)
0 = normal (27.5-40.7)
1 = mildly fast (40.8-44.1)
2 = moderately fast (44.2-47.5)
3 = severely fast (47.6-50.9)
4 = profoundly fast (51 이상)

-노년층: 27.8±2.9 (단위: 회)

-4 = profoundly slow (5.0 이하)
-3 = severely slow (5.1-10.7)
-2 = moderately slow (10.8-16.4)
-1 = mildly slow (16.5-22.1)
0 = normal (22.2-33.4)
1 = mildly fast (33.5-36.3)
2 = moderately fast (36.4-39.2)
3 = severely fast (39.3-42.1)
4 = profoundly fast (42.2 이상)

· 규칙성

0 = normal (정상)
1 = mild(약간 불규칙하나 음절 간 뭉침이 없음)
2 = moderate(자주 불규칙하나 음절 간 뭉침이 없음)
3 = severe (대부분 불규칙하며 음절 간 뭉침이 간헐적으로 관찰됨)
4 = profound (대부분 불규칙하며 음절 간 구분이 이 안 됨)

14-15. AMR /커/ 반복횟수 및 규칙성

· 반복횟수

-청년층: 34.1±3.9 (단위: 회)
-4 = profoundly slow (7.3 이하)
-3 = severely slow (7.4-14.0)
-2 = moderately slow (14.1-20.7)
-1 = mildly slow (20.8-27.4)
0 = normal (27.5-40.7)
1 = mildly fast (40.8-44.1)
2 = moderately fast (44.2-47.5)
3 = severely fast (47.6-50.9)
4 = profoundly fast (51 이상)

-노년층: 27.1±2.5 (단위: 회)
-4 = profoundly slow (4.3 이하)
-3 = severely slow (4.4-10)
-2 = moderately slow (10.1-15.7)
-1 = mildly slow (15.8-21.4)
0 = normal (21.5-32.7)
1 = mildly fast (32.8-35.6)
2 = moderately fast (35.7-38.5)
3 = severely fast (38.6-41.4)
4 = profoundly fast (41.5 이상)

· 규칙성

0 = normal (정상)
1 = mild (약간 불규칙하나 음절 간 뭉침이 없음)
2 = moderate (자주 불규칙하나 음절 간 뭉침이 없음)
3 = severe (대부분 불규칙하며 음절 간 뭉침이 간헐적으로 관찰됨)
4 = profound (대부분 불규칙하며 음절 간 구분이 안 됨)

16-17. SMR (조음 정확성을 유지하면서 5 초 동안 가능한 빨리 규칙적으로 /퍼터커/를 반복)[4]

· 반복횟수

-청년층: 12.0±1.1 (단위: 회)
-4 = profoundly slow (2.8 이하)
-3 = severely slow (2.9-5.1)
-2 = moderately slow (5.2-7.4)
-1 = mildly slow (7.5-9.7)
0 = normal (9.8-14.2)
1 = mildly fast (14.3-16.5)
2 = moderately fast (16.6-18.8)
3 = severely fast (18.9-21.1)
4 = profoundly fast (21.2 이상)

-노년층: 9.9±1.0 (단위: 회)
-4 = profoundly slow (1.5 이하)
-3 = severely slow (1.6-3.6)
-2 = moderately slow (3.7-5.7)
-1 = mildly slow (5.8-7.8)
0 = normal (7.9-11.9)
1 = mildly fast (12.0-14.0)
2 = moderately fast (14.1-16.1)
3 = severely fast (16.2-18.2)
4 = profoundly fast (18.3 이상)

· 규칙성

0 = normal (정상)
1 = mild(약간 불규칙하나 음절 간 뭉침이 없음)
2 = moderate(자주 불규칙하나 음절 간 뭉침이 없음)
3 = severe (대부분 불규칙하며 음절 간 뭉침이 간헐적으로 관찰됨)
4 = profound (대부분 불규칙하며 음절 간 구분이 안 됨)

18. bradyphemia (전체 발화 중 말속도가 점차로 느려지는 발화의 비율)[5]

0 = normal (정상)
1 = mild (25% 이하)
2 = moderate (26-50%)

3 = severe (51-75%)
4 = profound (76% 이상)

19. tachyphemia (전체 발화 중 말속도가 점차로 빨라지는 발화의 비율)

0 = normal (정상)
1 = mild (25% 이하)
2 = moderate (26-50%)
3 = severe (51-75%)
4 = profound (76% 이상)

20. conversation rate (정상 성인: 5.2±0.96, 단위: 초당 음절수)[6]

-4 = profoundly slow (1.32 이하)
-3 = severely slow (1.33-2.29)
-2 = moderately slow (2.30-3.26)
-1 = mildly slow (3.27-4.23)
0 = normal (4.24-6.16)
1 = mildly fast (6.17-7.13)
2 = moderately fast (7.14-8.10)
3 = severely fast (8.11-9.07)
4= profoundly fast (9.08 이상)

21. inappropriate silence

0 = normal (정상)
1 = mild (2 초 이하)
2 = moderate (3-4 초)
3 = severe (4-5 초)
4 = profound (5 초 이상)

22. stuttering (전체 자발화 중 반복, 연장, 막힘 등의 말더듬 증상 비율)[7]

0 = normal (2% 이하)
1 = mild (3-7%)
2 = moderate (8-15%)
3 = severe (16-25%)
4 = profound (26% 이상)

23. palilalia[8]

0 = normal (정상)
1 = mild (전체 자발화 중 2% 이하로 관찰됨)
2 = moderate (전체 자발화 중 3-7%, 읽기 과제에서 가끔 관찰됨)
3 = severe (전체 자발화 중 8-15%, 읽기 과제에서 자주 관찰됨)
4 = profound (전체 자발화 중 16% 이상, 읽기 과제에서 항상 관찰됨)

24. precise consonant (자발화의 자음정확도)

0 = normal (96% 이상)
1 = mild (66-95%)
2 = moderate (36-65%)
3 = severe (6-35%)
4 = profound (5% 이하)

25. speech intelligibility (자발화 수준의 명료도)

0 = normal (96% 이상)
1 = mild (66-95%)
2 = moderate (36-65%)
3 = severe (6-35%)
4 = profound (5% 이하)

26. speech awareness (환자의 보고)

0 = normal (발병 전과 차이가 없으며, 일상생활에서 전혀 불편함이 없음)
1 = mild (발병 전과 약간의 차이가 있으며, 일상생활의 25% 이하에서 불편함을 느낌)
2 = moderate (발병 전과 뚜렷한 차이가 있으며, 일상생활의 26-50%에서 불편함을 느낌)
3 = severe (발병 전과 뚜렷한 차이가 있으며, 일상생활의 51-75%에서 불편함을 느낌)
4 = profound (발병 전과 전혀 다르다고 느끼며, 일상생활의 76% 이상에서 불편함을 느낌)

---

1) MPT: Maximum Phonation Time
1) 2) 4) 김은정(2003). 『정상 청년층과 노년층의 최대발성시간 및 조음교대운동속도 비교』. 연세대학교 대학원 석사학위 논문.
AMR: Alternate Motion Rate; SMR: Sequential Motion Rate.
3) 연령별 1SD 간격은 /퍼/, /터/, /커/의 SD평균값을 사용함; profoundly slow에서 normal까지는 2SD를, mildly fast에서 profoundly fast까지는 1SD를 기준으로 함.
5) 본 연구자들에 의해 제안된 용어임.
6) "이영미, 김향희. 2001. 대화와 그림설명을 통한 한국 성인 발화의 비교분석. 『언어청각장애연구』, 6, 40-52. 서울: 한국언어청각임상학회."에 근거한 기준임.
7) "M. Wingate, 1981. 『Stuttering Theory and Treatment』. New York: Irvington Publishing Co."에 근거한 기준임.
8) normal 및 mild인 경우, 읽기과제에서 나타나는 동어반복은 자가수정으로 간주함.