

마비말장애 환자의 말명료도
측정을 위한 등간척도법 비교

연세대학교 대학원
언어병리학 협동과정
박 혜 진

마비말장애 환자의 말명료도
측정을 위한 등간척도법 비교

지도 김 향 희 교수

이 논문을 석사 학위논문으로 제출함

2007년 6월 일

연세대학교 대학원
언어병리학 협동과정
박 혜 진

박혜진의 석사 학위논문을 인준함

심사위원 _____ 인

심사위원 _____ 인

심사위원 _____ 인

연세대학교 대학원

2007년 7월 일

감사의 글

이 한 편의 논문에는 많은 분들의 노고가 담겨있습니다. 그 모든 분들의 맘과 정성으로 이루어진 논문이기 때문에 이 논문은 저만의 것이 아니며, 따라서 이후 더 값진 연구 자료로 쓰일 수 있기를 바랍니다.

지난 2년간 저희를 지도해주시며 본 연구의 기틀을 잡아주시고 마지막까지 열성으로 꼼꼼하게 지도해주신 김향희 교수님, 연구 방법을 세심하게 지도해주신 박은숙 교수님, 통계 부분에서 많은 지도를 해주신 김덕용 교수님께 진심으로 감사드립니다.

본 연구에서 가장 힘들었던 부분은 피험자와 평가자 선생님들을 구하는 일이었습니다. 바쁘신 중에 귀찮고 힘든 일이었는데도 선후배라는 이름 하나에 선뜻 부탁을 들어주셨던 많은 선생님들께 진심으로 감사드립니다. 김현승 선생님, 박지은 선생님, 방영임 선생님, 백여정 선생님, 서미경 선생님, 윤지혜 선생님, 이경하 선생님, 이미숙 선생님, 이영미 선생님, 이혜란 선생님의 도움이 없었다면 본 논문은 완성되지 못했을 것입니다. 또한 바쁜 실습 중에도 싫은 내색 없이 평가에 임해준 사랑스런 13기 후배들과 윤희연니, 지원이에게 깊은 감사의 마음을 전합니다. 막강한 동기애를 자랑하며 지난 2년간 동고동락하면서 많은 추억을 쌓았던 12기 동기들에게도 고마운 마음을 전하며 모두 훌륭한 언어치료사로 거듭날 수 있기를 기원합니다.

이 외에도 통계문제로 고민할 때 많은 조언을 해주신 남정모 교수님과 황선호 오빠에게 감사드리며, 논문 진행에 어려움이 있을 때마다 물심양면으로 도와주시며 관심 가져주신 여러 선배 선생님들께 감사의 말씀을 드립니다. 특히 제게 이 길을 걷게 해주시고, 늘 용기를 주시며 응원해주는 이정택 교수님께도 진심으로 감사의 말씀을 드립니다.

마지막으로 늘 마음으로 지켜봐주고 든든한 버팀목이 되어주는 가족들에게 사랑의 마음을 전합니다. 또 이제 새로운 가족이 될 부모님과 7년간 한결같은 사랑과 믿음으로 곁을 지켜주었던 그에게도 감사와 사랑의 마음을 전합니다.

연구자 씬

차 례

표 차례	iv
국문 요약	1
I. 서론	3
1. 이론적 배경	3
가. 마비말장애 환자의 말명료도 평가	3
나. 말명료도 평가에 영향을 미치는 요인	5
다. 연구의 필요성	8
2. 연구 문제	10
II. 연구 대상 및 방법	12
1. 연구 대상	12
가. 말명료도 평가자 정보	12
나. 말명료도 검사의 피검자 정보	14
2. 연구 방법	19
가. 자료 수집	19
나. 평가 과제	20
다. 평가 과정	21
라. 통계 분석	22

III. 결 과	23
1. 언어치료전문가 집단과 대학원생 집단 간 등간척도법의 공인타당도 비교	23
2. 언어치료전문가 집단과 대학원생 집단 간 등간척도법의 신뢰도 비교	24
가. 평가자내 신뢰도	24
(1) 3점 등간척도법	24
(2) 5점 등간척도법	25
(3) 7점 등간척도법	26
나. 평가자간 신뢰도	27
(1) 3점 등간척도법	27
① 언어치료전문가 집단	27
② 대학원생 집단	28
(2) 5점 등간척도법	29
① 언어치료전문가 집단	29
② 대학원생 집단	30
(3) 7점 등간척도법	31
① 언어치료전문가 집단	31
② 대학원생 집단	32
3. 세 가지 등간척도법 비교	33
IV. 고 찰	34
V. 결 론	38

참고 문헌	40
부록	45
영문 요약	51

표 차례

표 1. 말명료도 평가자 정보	13
표 2. 말명료도 검사의 피검자 정보	16
표 3. 피검자의 말명료도 (말명료도 평균 순 배열)	18
표 4. 두 집단의 세 가지 등간척도법의 공인타당도	23
표 5. 3점 등간척도법의 평가자내 신뢰도	24
표 6. 5점 등간척도법의 평가자내 신뢰도	25
표 7. 7점 등간척도법의 평가자내 신뢰도	26
표 8. 3점 등간척도법에서 언어치료전문가 집단의 평가자간 신뢰도	27
표 9. 3점 등간척도법에서 대학원생 집단의 평가자간 신뢰도	28
표10. 5점 등간척도법에서 언어치료전문가 집단의 평가자간 신뢰도	29
표11. 5점 등간척도법에서 대학원생 집단의 평가자간 신뢰도	30
표12. 7점 등간척도법에서 언어치료전문가 집단의 평가자간 신뢰도	31
표13. 7점 등간척도법에서 대학원생 집단의 평가자간 신뢰도	32
표14. 세 가지 등간척도법에 대한 두 평가자 집단 간 타당도와 신뢰도 비교	33

국 문 요 약

마비말장애 환자의 말명료도 측정을 위한 등간척도법 비교

말명료도란 환자의 말을 상대방이 의미적으로 알아들을 수 있는 정도를 말한다. 말명료도를 평가함으로써 말의 하부체계들이 독립적이면서 상호적으로 기능하는지, 실제로 환자가 타인과 어느 정도의 대화가 이루어질 수 있는지를 알아볼 수 있다. 말명료도 평가에 가장 널리 사용되고 있는 청지각적 평가방법은 등간척도법으로서, 청자가 각 자극을 듣고 눈금이 그려진 연속적인 선 위에 환자의 상태에 해당하는 수를 표시하는 방법이다. 여러 연구자가 각자 사용하는 등간척도 점수 종류는 다양하지만 일정한 기준 없이 척도 점수를 사용하고 있으며, 어떤 척도 점수가 가장 타당도와 신뢰도가 높은지에 대한 연구가 이루어지지 않았다.

이에, 본 연구에서는 마비말장애 환자의 말명료도를 3점, 5점 및 7점의 세 가지 등간척도법으로 평가하여 각 척도 점수에 대한 타당도와 신뢰도를 알아보고자 하였다. 또한 언어치료 경험의 여부에 따라 평가 집단을 둘로 나누어 언어치료 경험이 있는 평가자 집단과 경험이 없는 평가자 집단 사이에 평가 척도에 따라 수행력이 차이가 있는지 알아보고자 하였다.

본 연구를 위해 17명의 평가자를 모집하였는데, 7명은 1년 이상의 언어치료 경험이 있는 언어치료전문가들이며, 10명은 치료 경험이 없는 대학원생들이었다. 이들은 미리 녹음된 25명의 마비말장애 환자들의 문단읽기 발화 샘플을 들으면서 3, 5, 7점의 등간척도법을 사용해 말명료도를 평가하였다. 등간척도평가 결과는 일음절 낱말대조 명료도 검사 결과와 비교하여 Pearson 상관분석을 통해 공인타당도 검정을 하였다. 또한 평가자내 신뢰도와 평가자간 신뢰도를 알아보기 위해 Pearson 상관분석을 실시하였다.

본 연구의 결과는 다음과 같다.

첫째, 공인타당도의 경우 세 가지 등간척도법에 대해 언어치료전문가와 대학원생 집단이 모두 유의한 타당도를 보였다.

둘째, 평가자내 신뢰도의 경우, 3점 척도를 사용했을 때, 언어치료전문가와 대학원생 집단에서 각각 43%와 30%의 평가자만이 유의한 신뢰도를 보였다. 반면, 5점 척도에서는 언어치료전문가와 대학원생 집단의 평가자 모두가 유의한 신뢰도를 보였다. 7점 척도에서는 언어치료전문가 집단 평가자는 모두 유의한 신뢰도를 보인 반면, 대학원생 집단은 60%의 평가자만 유의한 신뢰도를 보였다.

셋째, 평가자간 신뢰도의 경우에는 두 집단이 공통적인 양상을 보였다. 3점 척도의 경우 두 집단에서 각각 두 쌍의 평가자들이 유의하지 않은 신뢰도를 보였으나, 5점 및 7점 척도에서는 두 집단이 모두 유의한 신뢰도를 보였다.

본 결과들을 종합해보면, 언어치료전문가 집단에서는 5점 척도와 7점 척도가, 대학원생 집단에서는 5점 척도만이 타당도 및 신뢰도를 가진 평가법으로 나타났다. 따라서 치료경험의 여부와 상관없이 타당도와 신뢰도가 높은 평가방법은 5점 척도라고 할 수 있다. 그러나 평가자의 치료경험이 많아지면 평가자의 신뢰도가 높아지므로 7점 척도 역시 타당도와 신뢰도가 높은 평가방법이 될 수 있을 것이다. 이러한 결과는 5점 척도가 환자의 상태를 충분히 설명할 정도로 등간이 세분화되어 있으며 평가 신뢰도에 방해를 주지 않는 좋은 평가법이라고 인식되어온 이전의 주장을 지지해주며, 평가자의 치료경험이 증가할수록 평가자의 신뢰도가 높아진다는 이전의 연구를 뒷받침한다. 본 연구의 의의는 이러한 주장에 대한 구체적 근거를 제시해 주었다는 점이다. 이에 따라, 향후 여러 연구와 치료에서 다양하게 사용되었던 등간척도 점수를 통일해서 사용하는데 도움이 될 수 있을 것이다.

핵심이 되는 말: 마비말장애, 말명료도, 청지각적 평가, 등간척도법

마비말장애 환자의 말명료도 측정을 위한 등간척도법 비교

<지도교수 김 향 희>

연세대학교 대학원 언어병리학 협동과정

박 혜 진

I. 서 론

1. 이론적 배경

가. 마비말장애 환자의 말명료도 평가

말을 하기 위해서는 호흡기관, 발성기관, 공명기관, 조음기관 등의 말하부기관이 필요하다. 이런 하부기관들은 중추신경계와 말초신경계의 지배를 받게 되는데, 이 때 중추 및 말초신경계에 손상이 생기면 근육이 마비되거나 약해지며, 속도가 저하되고 협응이 어려워지게 된다. 이런 말운동장

애를 마비말장애(dysarthria)라고 한다.¹ 마비말장애의 특징은 주어진 과제 및 상황에 따라 말 오류 형태들이 대체적으로 일관되게 관찰된다는 것이다. 또한, 마비말장애는 기능이 저하되는 말 하부기관이 무엇인가에 따라 독특한 특징이 관찰된다.¹ 예를 들어 조음기관의 경우 입술, 턱, 혀 등이 포함되는데, 이런 조음에 관련된 구조적인 기관들이 기능하여 조음이라는 하부체계를 구성하게 된다. 하부체계는 각 하부기관에서 이루어지는 말 영역의 기능적 관점이다. 따라서 이러한 오류의 일관성과 하부체계의 특성에 비추어 마비말장애 여부에 대한 정확한 평가가 이루어져야할 것이다. 또한 일단 마비말장애로 진단이 내려지면 어느 정도의 중증도를 가지고 있는지, 말의 하부체계 중 어떤 부분에서 가장 문제를 보이는지 등을 정확히 평가해야 그에 적절한 치료 계획이 이루어질 수 있을 것이다.

그렇다면 마비말장애의 어떤 영역을 평가해야할 것인가? 많은 선행연구에서 연구자들은 말의 하부체계를 중심으로 평가 영역을 세분화하였다. Mayo Clinic의 Darley, Aronson과 Brown(이하 DAB 연구)^{2,3}은 마비말장애 평가영역을 말의 하부체계를 중심으로 여덟 가지의 영역으로 나누고, 이를 바탕으로 38개의 청지각적 요소를 평가하였다. Duffy¹는 DAB의 연구에 10가지의 말 특성을 추가하여 총 48개의 청지각적 요소를 제시하였다. Cahill 등⁴ 역시 TBI(Traumatic Brain Injury) 아동의 말을 평가하면서 호흡, 발성, 공명, 조음, 운율의 하부체계를 중심으로 33개의 청지각적 요소를 평가하였다. 결국 다섯 가지 말의 하부체계는 각각 독립적인 평가영역이지만 궁극적으로 전반적인 말산출과 연관되어 있다. 말산출이 얼마만큼 정상적으로 이루어지는가에 대한 판단은 말명료도(speech intelligibility)를 측정함으로써 가능하다. 말명료도란 환자의 말을 대화의 상대방이 의미적으로 알아들을 수 있는 정도를 말한다.⁵ 이러한 말명료도가 정상적인 상태를 유지하기 위해서는 다섯 가지 하부체계의 독립적인 그리고 상호적인 기능이 정

상적으로 발휘되는 것이 중요하다. 이 때문에 말명료도를 평가하는 것이 중요하며, 실제 환자가 타인과 어느정도의 대화가 이루어질 수 있는지를 평가할 수 있는 직접적인 평가라고 할 수 있다. 치료적으로 말명료도 평가가 중요한 이유는 조음 뿐 아니라 호흡, 공명, 발성 등의 말 산출에 필요한 모든 요소들을 반영하는 측정치이기 때문이며,⁶ 명료도가 말 운동장애정도를 반영하는 길항운동이나 조음정확도와 같은 측정치들과 높은 상관관계를 나타내는 자료이기 때문이다.⁷ 또한 별다른 생리학적 지식이 없는 당사자와 가족들에게 쉽게 결과를 해석해줄 수 있다는 장점이 있다.⁸

나. 말명료도 평가에 영향을 미치는 요인

말명료도에 영향을 주는 요인은 일반적으로 화자, 청자, 검사도구의 내용 및 평가방법의 세 가지 측면으로 나뉘어진다.⁹ 화자와 관련된 요인은 산출된 말의 조음적 요인, 언어적 요인 등이 있다.¹⁰ 선행연구들에 의하면 조음적 요인과 말명료도는 높은 상관관계를 가지며, 말명료도에 가장 큰 영향을 주는 변수로 나타난다.^{11,12} 청자관련 요인으로는 화자와 평가자와의 친밀도, 경험 등을 들 수 있다. Osberger¹³는 청자가 마비말장애 환자의 말을 들어 본 경험이 있거나, 환자와 공통된 기억을 갖고 있을 때, 또 환자가 말하는 내용을 알고 있을 때 말명료도가 증가하다고 보았다. 검사도구와 관련해서는 검사도구의 내용과 평가방법이 말명료도 측정에 영향을 줄 수 있다. 즉 검사도구의 내용으로 어느 것을 사용했는지 등이 평가에 영향을 미칠 수 있을 것이다.¹⁴ 또한 평가방법에 따라서도 영향을 받을 수 있는데 평가방법에는 등간이나 비율척도를 사용하여 청자가 주관적으로 표시하는 방법과 화자의 말을 듣고 청자가 받아 적은 후 검사도구의 내용과 일치하는 낱말이나 음절의 퍼센트를 계산하는 방법이 있다.¹⁵

화자, 청자, 검사도구의 내용 및 평가방법의 세 가지 요인 중 실제로 언어치료전문가가 환자를 평가 할 때 신뢰도에 영향을 주게 되는 것은 검사도구와 관련된 요인과 치료 경험 정도이다. 대체로 치료 경험과 관련해서는 여러 연구에서 치료 경험이 높은 언어치료전문가의 신뢰도가 더 높다는 결과가 보고되고 있다.^{16,17} 따라서 치료경험이 많지 않은 평가자는 이 약점을 보완하기 위해 다른 방해 요소를 최대한 줄이는 노력이 필요할 것이다.

검사도구에 따른 요인 중 도구의 내용과 관련하여 어떤 발화내용으로 평가를 했는지에 따른 신뢰도의 차이에 대한 연구도 활발하게 진행되었다. Kempler와 Lancker¹⁸는 자발화, 따라말하기, 읽기, 노래 따라부르기, 스스로 노래부르기의 과제를 통해 신뢰도를 검증한 결과, 다른 과제에서는 대체로 80% 전후에서 말명료도가 평가되었으나 자발화 수준에서의 말명료도는 29%로 현저히 낮게 평가되었다. 이는 같은 환자의 말이라도 어떤 발화내용을 평가하느냐에 따라 상대가 알아들을 수 있는 정도에 차이가 있음을 보여준다. 일반적으로 언어능력을 평가하기 위하여 가장 타당도가 높은 검사방법은 자연스러운 대화 가운데서 발화를 수집하여 분석하는 것이지만, 다양한 음소들이 다양한 맥락에서 통제되어 산출되기가 어려운 실정이므로 언어평가 시 문단읽기과제를 채택하고 있다.⁸ 따라서 본 연구에서도 읽기 문단을 통해 발화 음소를 통제하여 말명료도를 평가하고자 한다. 그러나 읽기는 말하기에 비해 화자들이 더 정확하게 조음하는 경향이 있으므로, 일상적인 말하기와는 차이가 있음을 감안하여 해석할 필요가 있다.⁸

Whitehill 등¹⁹은 검사도구의 평가 방법을 비교하였다. 과비성에 대한 등간척도법(Equal-Appearing Interval scale, EAI)과 DME(Direct Magnitude Estimation) 척도법의 신뢰도를 비교한 연구로, DME 척도법을 다시 DME-M(Direct Magnitude Estimation with Modulus)과 DME-WM(Direct Magnitude Estimation without Modulus)으로 나누어 세 방법에 대한 신뢰

도를 측정하였다. DME-M은 정상 기준치가 제시된 경우이며, DME-WM은 기준치가 제시되지 않고 처음 들려준 샘플을 기준으로 삼아 비교 평가하는 방법이다. 이 연구에서 등간척도법과 DME-M, DME-WM의 세 평가 방법은 곡선그래프를 보이며 서로 상관관계를 보였지만, DME의 평가 방법이 더 신뢰도가 높은 평가방법이라고 보고하고 있다. 또한 Samar와 Metz,²⁰ Schiavetti,¹⁵ Kent²¹는 평가자가 환자의 말을 듣고 이해한 말의 단어 수나 음소 수를 전사하거나, 제시된 보기에서 골라 맞은 정도를 퍼센트로 계산하는 방법이 척도에 의한 방법보다 신뢰도가 높다고 보고하였다. 실제로 Hustad²²는 말명료도를 측정하는 청지각적 판단에 의한 평가 방법을 설명하면서 이렇게 단어 및 문장 테스트를 이용하여 퍼센트를 계산하는 방법을 객관적 측정법(Objective Measure of Intelligibility)으로 분류하였다.

그러나 실제 치료에서는 마비말장애 환자의 말명료도를 평가할 때 DME 평가 방법과 단어 및 문장 테스트를 이용한 퍼센트 계산법을 잘 사용하지 않는다. 그것은 DME 평가 방법을 사용할 경우 해당 성별, 연령에 맞는 기준 말 샘플을 준비하기가 번거로우며,²³ 치료에서 환자를 보고 바로 평가를 내려야하는 상황에서 매번 말 샘플을 듣고 그와 비교해 환자를 평가하기가 어렵기 때문이다. 단어 및 문장 테스트를 이용한 퍼센트 계산법도 평가 시간이 많이 소요되는 단점이 있다. 또한 음성의 질이나 운율같은 초분절적인 요소가 말 명료도에 영향을 미치는데 이를 전사하여 표기하는 것이 어려우며, 완전한 발화 내에서 손상 정도를 살피지 못하고 단어나 문장에 기초하여 측정된다는 점도 제한점으로 꼽힌다.²³ 특히 우리나라에는 표준화되어 출판된 검사도구가 없기 때문에 비공식적으로 자신이 단어 목록을 만들어 사용해야하는 어려움이 있다. 현재 우리나라에서 연구된 말명료도 평가를 위한 단어 평가도구는 윤미선 등²⁴의 청각장애 아동의 말명료

도를 위한 단어 목록과, 김수진²⁵의 일음절 낱말대조 명료도 평가방법 정도를 찾아볼 수 있으나 이들 역시 공식검사로 출판되지 않는 상황이다. 이 때문에 치료에서는 비교적 신뢰도가 낮더라도 간편하게 사용할 수 있는 등간척도법이나 퍼센트 측정법(percent estimation)이 자주 사용된다. 등간척도법과 퍼센트 측정법은 DME 평가방법과 함께 말명료도의 주관적 측정법(Subjective Measure of Intelligibility)으로 분류되는 평가방법이다.²²

다. 연구의 필요성

등간척도법은 말 영역 평가에서 가장 널리 사용되고 있는 방법이다.¹⁹ 등간척도법에 의하면 청자는 각 자극을 듣고 눈금이 그려진 연속적인 선위에 환자의 상태에 해당하는 수를 표시 하게 된다.¹⁵ 척도에 활용되는 숫자 간격으로는 5점 척도, 7점 척도, 9점 척도 등의 홀수 척도가 많이 사용되는데 이는 중앙점수를 활용할 수 있기 때문이다.²⁶ 이 때 대체로 가장 낮은 수인 0이나 1을 정상(normal) 수준으로 하고, 가장 장애가 심한 정도를 가장 높은 숫자로 표시한다.^{1,2,4} Duffy¹는 5점 점수 척도를 사용하면서 0점을 정상(normal), 1점을 약함(mild), 2점을 보통(moderate), 3점을 두드러짐(marked), 4점을 심함(severe)로 표시하였다. Zraick와 Liss²⁷의 연구에서도 비성의 음질을 측정하는 방법으로 등간척도법과 DME 측정법을 비교하였는데 이때 등간척도법의 경우 5점 척도를 사용하였다. Hoodin과 Gilbert²⁸도 파킨슨씨 병(Parkinson's Disease) 환자의 과대비성을 측정하기 위해 5점 척도를 사용하였다. 한편, 마비말장애 연구의 고전이라 할 수 있는 DAB의 연구²에서는 여러 유형의 말장애에 대한 평가항목에서 7점 척도를 사용하여 평가하였다. Ludlow와 Vassich²⁹도 역시 파킨슨 병 및 샤이-드레거증후군(Shy-Drager Syndrome) 환자들을 대상으로 한 연구에서 총 20개

의 청지각적 평가항목 중에 17개 항목들에 대해 7점 척도를 사용하였으며, Karnell 등³⁰와 Workinger와 Kent³¹의 경우도 구개열 환자와 뇌성마비 아동의 과대비성 평가 시 7점 척도를 사용하였다. 이외 Hathaway³²은 pharyngeal flap 수술을 한 구개열 환자 연구 시 15점 척도를 사용하기도 하였으며, Cahill 등⁴의 경우 TBI 아동의 말 평가 시 말명료도와 음도끊김 (pitch breaks), 음도수준(pitch level) 등을 각각 4점, 5점, 7점 척도로 나눠 평가하기도 하였다.

이렇게 여러 연구자가 저마다 사용하는 점수 척도가 매우 다양하지만, 각 점수 척도를 사용하는 이유를 밝힌 연구는 없었다. Duffy¹의 경우도 5점 등간척도법을 사용하면서, 일반적으로 연구에서 5점 점수척도가 자주 사용되며, 많은 연구자와의 평가에서 일치도가 높다고 설명하고 있지만 그 구체적인 근거는 제시하지 않았다. 또한 등간척도법은 주관적인 평가방법으로 평가자의 평가 경험정도가 주요한 요소로 작용할 것이므로 이에 대한 검증도 필요하였다. 따라서 본 연구에서는 평가자가 신뢰도에 영향을 미치는 요인인 치료 경험 여부에 상관없이 타당하고도 신뢰로운 등간척도법을 제시하고자 하였다.

2. 연구 문제

Stevens³³의 연구에 의하면 청자는 등간척도법을 사용할 때 많은 현상에 의해 방해 받는다고 하는데, 등간척도의 한 간격을 보다 작게 세분화하는 것 등이 방해 요인으로 작용한다고 하였다. 이것은 결국 검사 도구에 의한 방해 요인으로서 등간척도법의 척도 간격에 따라 신뢰도 검정에 차이가 나타날 수 있음을 말하는 것이다. 확률적으로 척도 간격의 수가 적을수록 평가자간 일치도가 높아지겠지만, 척도 간격이 너무 넓으면 정확성이 떨어질 염려가 있다. 또한 말명료도 평가에 영향을 미치는 청자요인으로서 평가자의 치료 경험 정도가 평가에 영향을 미치는 주요한 요소로 지적되므로, 경험의 차이에 의해 적절하다고 제안할 수 있는 평가방법이 차이를 보이는지 알아볼 필요가 있다.

본 연구에서는 마비말장애 환자의 발화를 수집하여 비교적 객관적인 방법으로 제시된 단어를 이용한 퍼센트 계산법에 의한 점수를 기준으로 제시하여 이를 3점, 5점, 7점의 세 가지 종류의 등간척도법과 비교하여 어떠한 상관관계를 갖는지 알아보려고 한다. 이를 통해 높은 상관관계를 갖는 검사도구가 타당도가 높은 검사방법이라는 결론을 제시할 수 있을 것이다. 또한 세 가지 종류의 등간척도법의 신뢰도는 각 평가자들의 평가자내 신뢰도와 평가자간 신뢰도를 비교하여 알아보려고 한다. 성태제³⁴에 의하면 두 평가방법의 상관계수가 0.6 이상이면 타당도 및 신뢰도가 높다고 할 수 있다고 하였다. 따라서 타당도와 신뢰도 계수가 모두 0.6 이상을 만족하는 평가방법을 말명료도에 적합한 평가방법이라고 제안할 수 있을 것이다. 이때 평가자로 마비말장애 환자에 대한 치료 경험이 있는 언어치료전문가 집단과 경험이 없는 대학원생 집단으로 구분하여, 경험 여부에 따라 타당도와 신뢰도가 높은 평가 방법에 차이가 있는지 알아보려고 한다.

본 연구의 연구문제는 다음과 같다.

- 1) 평가 치료 경험 유무에 따라 가장 높은 타당도 및 신뢰도를 보이는 등간척도법은 3점, 5점, 그리고 7점 척도법 중 어떤 것인가?

II. 연구 대상 및 방법

1. 연구 대상

가. 말명료도 평가자 정보

본 연구를 위해 평가자는 마비말장애 환자에 대한 경험이 있는 언어치료전문가 집단과 경험이 없는 대학원생 집단으로 구분하였다.(표 1) 언어치료전문가 집단은 모두 7명으로, 서울 및 경기 지역에 위치한 병원의 재활의학과 및 신경과에서 근무 중이었다. 이들은 1급 언어치료사 자격증을 소지하고 있으며, 일주일에 적어도 5시간 이상 마비말장애 환자의 평가 및 치료를 담당하고 있는 경력이 1년(총 240시간) 이상인 언어치료전문가로 제한을 두었다. 이 평가자군의 평균 연령(\pm 표준편차)은 32.29(\pm 3.451)세이며 치료경험 년 수는 평균(\pm 표준편차) 4.14(\pm 3.390)년이었다.

대학원생 집단은 언어병리학 대학원 석사과정 3학기 이상의 학생 10명을 대상으로 하였다. 이들은 한 학기 이상 마비말장애 수업을 수강한 경험이 있으나 마비말장애 환자에 대한 치료경험이 없는 평가자였다. 대학원생 집단은 모두 여자이며, 평균 연령(\pm 표준편차)은 25.90(\pm 1.454)세였다.

표 1. 말명료도 평가자 정보

평가자	성별	연령 (세)	언어 치료사 자격증	치료 경험 년 수	근무지 혹은 대학원학기	
언어 치료 전문 가	R01	여	36	1급	11	서울 재활의학과
	R02	여	33	1급	6	서울 신경과
	R03	여	27	1급	2	서울 재활의학과
	R04	여	30	1급	3	서울 재활의학과
	R05	여	31	1급	1	서울 재활의학과
	R06	여	37	1급	3	서울 신경과
	R07	여	32	1급	3	경기도 재활의학과
대 학 원 생	R08	여	28	없음	없음	5학기
	R09	여	29	없음	없음	5학기
	R10	여	25	없음	없음	4학기
	R11	여	25	없음	없음	3학기
	R12	여	25	없음	없음	3학기
	R13	여	25	없음	없음	3학기
	R14	여	25	없음	없음	3학기
	R15	여	26	없음	없음	3학기
	R16	여	25	없음	없음	3학기
	R17	여	26	없음	없음	3학기

나. 말명료도 검사의 피검자 정보

본 연구는 서울 및 경기 지역에 위치한 병원에서 언어치료전문가에 의해 마비말장애로 진단된 환자 25명을 대상으로 하였다. 진단내용은 표준화된 검사도구는 아니지만 마비말장애 유무를 감별할 수 있도록 고안된 최대모음연장 과제, 길항운동속도 과제, 문단읽기 과제 사용 등으로 담당 언어치료전문가에게 내용을 보고받았다.³⁵ 환자의 연령과 성별, 언어장애 원인과 정도는 제한을 두지 않고 가능한 다양하게 모집하였다.(표 2) 다양성에 대한 1차 평가는 담당 언어치료전문가의 평가에 의해 보고받았으며, 2차 평가는 각 환자의 일음절 낱말대조 명료도 평가 점수의 평균을 통해 살펴 보았다. 환자 25명의 일음절 낱말대조 명료도 평가 점수의 평균 점수(\pm 표준편차)는 74.86(\pm 13.269)이었다. 가장 점수가 낮은 환자 즉, 말명료도가 가장 저하되어 있는 환자의 평균(\pm 표준편차) 점수는 41.65(\pm 7.722)이었으며, 가장 높은 점수를 받아 말명료도가 가장 높은 환자의 평균(\pm 표준편차) 점수는 92.12(\pm 3.271)이었다. 분포를 살펴보면 60점 이하 4명, 60~70점 5명, 70~80점 6명, 80~90점 7명, 90점 이상 3명으로 다양하게 모집되어 있다. (표 3)은 말명료도 수준이 높은 환자부터 정렬해놓은 것이다.

다른 동반장애 유무에는 제한을 두지 않았지만, 읽기 수행이 가능해야 하므로 담당 언어치료전문가에 의해 문단 읽기가 가능하며, 시력손상이나 기타 무시증이 없다고 보고된 환자를 대상으로 하였다. 또한 실어증이 동반되는 경우 검사 절차를 이해하지 못하는 경우가 발생하지 않도록 한국판 웨스턴실어증검사(Paradise · Korean version-the Western Aphasia Battery: P · K-WAB)³⁶ 결과, 알아듣기 검사에서 40점 이상(총 200점)의 점수를 획득한 경우로 제한하였다. 40점은 P · K-WAB 검사를 통한 실어증 유형별 검사표 중 모든 언어 능력이 제한되어있다고 알려진 전반실어증(global

aphasia) 환자의 평균(40.30 ± 19.618)에 근거한 것이다.³⁶ 또한 15세 이하 환자의 경우 P·K-WAB 검사가 불가능하기 때문에 별도의 언어능력 검사를 통해 수용언어능력이 $-2SD$ 내에 속하는 아동을 대상으로 하였다. 모집된 말명료도 검사의 피검자는 여자는 6명, 남자는 19명이었으며, 연령은 5세부터 76세까지 다양하게 분포되었다. 평균(\pm 표준편차) 연령은 $49.68(\pm 22.234)$ 세였다. 교육년수는 무학이 3명, 1~6년이 6명, 7~11년이 3명, 12년 이상이 10명으로, 평균(\pm 표준편차) 교육년수는 $11.20(\pm 6.465)$ 년이었다. 발병 후 경과 기간은 평균(\pm 표준편차) $31.33(\pm 51.433)$ 일로 다양한 분포를 보였으며, 모두 오른손잡이였다. 명료도 평가자의 피검자에 대한 친숙도를 통제하기 위해, 평가자로 모집한 언어치료전문가가 치료 중인 환자는 모집 대상에 포함하지 않았다.

표 2. 말명료도 검사의 피검자 정보

피 검 자	성 별	연 령 (세)	교육 년수	POT (개월)	진단명	언어 이해 능력	말- 언어 장애	말관련 증상	손 잡 이
P01	남	12	6	18.17	TBI	79 %ile ¹	dys.	reduced loudness	Rt.
P02	남	54	16	0.77	Lt. MCA inf.	200 ²	dys.	rough voice	Rt.
P03	남	72	16	48.87	Lt. MCA inf.	183 ²	dys.	strained, strangle voice	Rt.
P04	여	5	6	24.27	skull defect	74 %ile ¹	dys.	defective articulation	Rt.
P05	남	72	6	14.17	Korsakoff's syn.	192 ²	dys. aph.	vocal fry	Rt.
P06	여	54	9	214.83	hypoxic brain damage	194 ²	dys.	strained, strangle voice	Rt.
P07	남	76	18	5.0	Lt. MCA inf.	79 ²	dys.	rough voice	Rt.
P08	남	26	14	2.87	Lt. MCA inf.	123 ²	dys.	strained, strangle voice	Rt.
P09	남	45	16	120.0	brain tumor	100 ²	dys.	strained, strangle voice	Rt.
P10	남	71	18	0.43	Lt. MCA inf.	96 ²	dys.	strained voice	Rt.
P11	남	67	16	121.0	PD	136 ²	dys.	hypophonia, harsh voice	Rt.
P12	여	46	0	2.73	Rt. ACA inf.	102 ²	dys	breathy voice	Rt.
P13	여	75	6	28.33	pontine inf.	76 ²	dys.	hoarse, breathy voice	Rt.
P14	남	43	16	1.06	Lt. BG ICH	64 ²	dys. aph.	reduced loudness	Rt.
P15	남	10	2	73.67	Moyamoya dis.	89 %ile ¹	dys.	hypernasal	Rt.
P16	남	42	16	1.13	Lt. medullary inf.	200 ²	dys.	monopitch	Rt.
P17	여	68	0	12.63	PD	200 ²	dys.	harsh, hoarse voice	Rt.

P18	남	49	16	0.23	brain tumor	158 ²	dys.	hoase, breathy voice	Rt.
P19	남	65	0	24.37	Rt. ACA inf.	180 ²	dys.	harsh, rough voice	Rt.
P20	남	55	16	12.73	Lt. BG ICH	144 ²	dys. aph.	harsh, rough voice	Rt.
P21	남	60	16	14.03	Lt. ACA inf.	153 ²	dys.	harsh, rough voice	Rt.
P22	남	36	18	1.0	Lt. BG ICH	100 ²	dys.	breathy voice	Rt.
P23	남	78	18	1.0	Lt. CR inf.	76 ²	dys.	rough voice	Rt.
P24	남	13	6	39.47	TBI	99 %ile ¹	dys.	imprecise consonant	Rt.
P25	여	48	9	0.37	Lt. BG ICH	187 ²	dys.	imprecise consonant	Rt.

¹ 15세 이하 아동의 경우 각 연령에 맞는 수용언어 검사도구로 검사한 후 수용언어능력의 백분위수를 기록하였음.

² 15세 이상의 성인의 경우 P·K-WAB 검사를 실시하여 AQ의 이해하기능력의 점수를 기록하였음(200점 만점).

POT : Post Onset Time, AQ : Aphasia Quotient, dys : dysarthria, aph : aphasia,

Lt. : Left, Rt. : Right, inf. : infarction, syn. : syndrome, dis. : disease,

MCA : Middle Cerebral Artery, ACA : Anterior Cerebral Artery,

ICH : Intra-Cerebral Hemorrhage, TBI : Traumatic Brain Injury,

PD : Parkinson's Disease, BG : Basal Ganglia, CR : Corona Radiata

표 3. 피검자의 말명료도 (말명료도 평균 순 배열)

	평균	표준편차
P16	94.12	3.350
P21	90.59	4.321
P24	90.02	7.385
P19	86.53	7.063
P25	86.47	3.222
P05	85.71	5.373
P23	85.29	4.974
P10	81.53	3.989
P11	81.35	5.278
P12	80.88	4.530
P20	79.76	7.815
P14	79.06	6.663
P15	76.41	5.562
P02	75.76	4.317
P18	74.18	4.420
P06	71.59	5.624
P22	69.53	7.372
P09	68.47	4.801
P04	68.24	6.408
P17	66.76	4.372
P01	65.53	4.677
P07	59.07	8.492
P03	58.57	10.101
P08	54.29	3.201
P13	41.65	7.310

2. 연구 방법

가. 자료 수집

각 환자에게 수집한 발화는 두 가지 종류이다. 먼저 객관적인 평가 척도를 얻기 위해 사용될 단어 테스트를 이용한 퍼센트 계산법으로, 단어는 김수진 등²⁵의 일음절 낱말대조 명료도 평가법의 단어 목록을 사용하였다. 전체 목록 160개를 모두 사용하기에는 환자와 평가자에게 모두 부담이 되므로 전체대조 문항으로 제시된 98개의 문항을 사용하였다. 전체 대조 문항만으로도 우리말 자음의 모든 장소 대조가 포함되므로 충분한 검사가 될 수 있다고 보았다. 전체대조 문항은 두 개의 낱말을 대조시킨 86개의 대조와 세 개의 낱말을 대조시킨 12개의 문항으로 구성되어 있다. 자세한 문항 내용은 (부록-1)과 같다. 환자는 연구자의 지시에 따라 제시된 카드의 단어를 읽도록 하였다. 평가자의 학습효과를 최소화하기 위해 각 환자에게 제시되는 문항의 순서는 임의대로 섞어서 제시하였다.

다음 본 평가에 해당하는 문단 수준의 읽기발화를 수집하였다. 문단은 성인 표준화 문구인 ‘가을’ 문단³⁷을 채택하였다. 발화 수집에서 문단읽기를 선택한 것은 자발화를 수집하였을 경우 환자와 발화 채취자의 대화 내용이 평가자가 발화를 이해하는데 영향을 줄 수 있기 때문이다. 또한 각 환자들의 특정 이야기가 평가자의 기억에 남아 학습효과를 가져올 수 있기 때문에 이를 배제하고자 하였다. 문단읽기 중에서 새로운 문단이 아닌 평가자들에게 익숙한 ‘가을’ 문단을 선택한 것은 두 가지 이유에서이다. 첫째는 ‘가을’ 문단은 현재 치료현장에서 마비말장애 환자의 언어를 평가할 때 가장 많이 사용하고 있는 문단으로서 본 연구를 통해 언어치료전문가들이 치료에서 제각기 사용하고 있는 평가 척도와 비교할 수 있게 함이다. 둘째는

앞서 언급했듯이 언어치료전문가들이 ‘가을’ 문단을 통해 환자들의 언어평가를 진행해오면서 문단의 내용이 매우 익숙해져있기 때문이다. 이러한 사실은 환자의 말명료도가 자발화보다 높게 측정될 수 있다는 약점을 갖게 되는 것이 사실이다. 그러나 이미 문단의 내용에 익숙해져 있어 오히려 학습효과는 최소화될 수 있다. 본 연구에서 한 환자에 대한 평가는 총 4회 이루어지게 된다. 새로운 내용의 문단을 평가자가 접했을 때 처음에는 문단의 내용을 이해하지 못해 낮은 점수를 주었다가 네 번의 반복을 거치는 동안 문단의 내용에 익숙해지면서 이해 정도가 증가할수 있다는 것을 우려하여 이를 방지하고자 하였다.

녹음은 소음이 적은 조용한 치료실에서 진행되었으며, 피검자가 제시된 읽기자료를 읽는 동안 음성은 Samsung Yepp MP3 YP-U2에 수집하였다. 마이크와 입술의 발화 방향이 이루는 각도는 왼쪽 아래 30도 정도로 유지하였다. 입술과 마이크로폰의 거리는 10~15cm로 일정하게 유지하도록 하였으며 샘플은 입력 단자를 이용하여 음성 파일로 저장하였다.

나. 평가과제

객관적 평가를 위한 일음절 낱말대조 명료도 평가는 평가자가 미리 녹음된 각 환자의 98개의 단어 발화를 듣고, 주어진 3개 혹은 5개의 보기 중에서 들은 단어를 고르도록 하는 선택형으로 제시되었다. 평가자는 각 단어를 한번만 듣고 제시된 보기 중에서 답을 골라 동그라미 표시를 하도록 하였다. 각 문항마다 5초간의 시간이 주어지며, 전체 98문항의 일음절 낱말대조 명료도 평가에 약 10분 정도가 소요되었다.

‘가을’ 문단에 대한 말명료도 평가를 위해, 환자가 읽은 전체 문단 중 가운데 부분의 네 문장만을 녹음하여 평가자에게 들려주었다. 환자가 문단

을 읽기 시작하면서 문단 읽기 과제에 익숙해지는 시간을 고려하여 처음 부분은 자극에서 제외하였으며, 문단 마지막 부분에서는 환자가 읽기에 지쳐 말명료도가 점차 떨어질 것을 우려하여 제외하였다.

‘가을’ 문단에 대한 등간척도법 점수에 대한 평가의 경우 3점 척도, 5점 척도, 7점 척도의 세 가지 종류의 등간척도 평가를 시행하였다. 서론에서 제시한 여러 연구에 의하면, 말명료도 평가에서는 7점 이상의 점수 척도를 잘 사용하지 않으며, 점수체계를 7점 이상 세분화하여 사용하는 경우는 주로 비성(nasality)이나 음도(loudness) 평가처럼 중앙 점수를 중심으로 과다(hyper-)와 과소(hypo-)를 모두 평가하고자 할 때 사용하였다.^{4,32}

본 연구에서 사용하는 각 척도는 낮은 점수일수록 말명료도가 높은 정상(normal)에 해당하는 수준이며, 높은 점수일수록 말명료도가 낮은 심한(severe) 수준의 마비말장애를 가리킨다. 즉, 각 척도의 0점은 정상(normal) 수준이며 각 척도의 가장 높은 숫자(2, 4, 6점)는 가장 심함(severe) 정도이다. 각 평가 항목마다 10초간의 시간을 주어 평가를 할 수 있도록 하였다.³⁸ 평가자는 환자의 발화를 모두 들은 후에 점수를 기입하도록 하며, 한번 평가된 점수는 다음 발화가 제시된 후에는 정정하지 않도록 지시하였다.

다. 평가 과정

17명의 평가자는 녹음된 25명의 마비말장애 환자의 발화를 듣고 평가척도에 따라 말명료도를 평가하였다. 평가자간 합의(consensus rating)는 허용되지 않았으며 별도의 청지각 훈련 역시 실시하지 않았다. 그러나 일음절 낱말대조 명료도 평가법의 경우 검사 방법이 평가자에게 낯설어 오류를 범할 수 있으므로 두 문항을 예로 제시하여 검사 방법을 숙지하도록 하였다. 평가자는 평가 전에 본 연구의 목적과 절차에 대해 설명을 들었으며,

평가 분야인 말명료도의 정의와 특징에 대해 충분히 이해할 수 있었다.

모든 평가는 4회에 걸쳐 이루어졌으며 각 회마다 일음절 낱말대조 명료도 평가 5명 분량과 척도평가 한 가지 방법이 포함되었다. 즉, 첫째 날에는 객관적 측정을 위한 일음절 낱말대조 명료도 평가 5명 분량과 3점 척도 평가가 함께 이루어지고 이후 24시간 간격²⁷으로 일음절 낱말대조 명료도 평가 5명분과 함께 5점 척도, 7점 척도가 각각 순서대로 이루어졌다. 4회째에 실시되는 재평가에서는 동일한 환자의 발화 샘플을 임의로 섞어 1주 후에 실시하였다. 각 평가자가 25명의 환자에 대한 평가를 모두 시행하는데 약 4시간 정도의 시간이 소요되었다.

라. 통계 분석

네 가지 평가 방법에 대한 공인타당도를 살펴보기 위해 객관적 평가방법인 일음절 낱말대조 명료도 평가방법과의 상관계수를 측정하였다. 두 가지 요인에 대한 상관관계는 Pearson 상관분석을 실시하였다. 평가자내 신뢰도(intra-judge reliability) 검정과 평가자간 신뢰도(inter-judge reliability) 검정을 위해 역시 각 값들 간의 상관계수를 측정하였으며, Pearson 상관분석을 실시하였다. 통계 프로그램은 SPSS(version 12.0)을 사용하였으며 통계학적 검정에 대한 유의수준은 0.05로 하였다.

III. 결 과

1. 언어치료전문가 집단과 대학원생 집단 간 등간척도법의 타당도 비교

언어치료전문가와 대학원생이 각 3가지 평가방법을 통해 환자의 말명료도를 평가했을 때, 모두 일음절 낱말대조 명료도 평가와 유의한 상관관계를 보이는 것으로 나타났다.(표 4) 세 가지 등간척도법이 모두 $p < 0.001$ 로 유의한 상관관계를 보였다. 이때, 언어치료전문가 집단의 경우 3점 척도의 상관계수가 -0.4 이상으로 상관관계가 있으며, 5점 척도와 7점 척도의 경우에는 상관계수가 -0.6 이상으로 상관관계가 높았다.³³ 언어치료전문가 집단의 경우 등간의 수가 증가할수록 상관관계가 점차 증가하였다. 대학원생 집단에서는 모든 점수 척도가 -0.6 이상으로 높은 상관관계를 보였다. 그러나 3점 척도와 7점 척도보다 5점 척도의 상관관계가 약간 더 높은 것을 볼 수 있다.

표 4. 두 집단의 세 가지 등간척도법의 공인타당도 (상관계수 r)¹

	언어치료전문가	대학원생
3점 척도	-0.577^{***}	-0.608^{***}
5점 척도	-0.633^{***}	-0.632^{***}
7점 척도	-0.642^{***}	-0.600^{***}

¹ 상관계수의 -값은 일음절 낱말대조 명료도 검사의 경우 점수가 높을수록 말명료도가 증가하지만, 등간척도법의 경우에는 점수가 높을수록 말명료도가 감소하므로 이러한 역의 상관관계를 의미함.

*** $p < 0.001$

2. 언어치료전문가 집단과 대학원생 집단 간 등간척도법의 신뢰도 비교

가. 평가자내 신뢰도

(1) 3점 등간척도법

3점 등간척도법으로 말명료도를 평가했을 때, 언어치료전문가와 대학원생 집단에서 각각 평가자내 신뢰도가 유의하지 않은 평가자가 관찰되었다. 언어치료전문가 집단에서는 7명 중 3명(43%)이, 대학원생 군에서는 10명 중 3명(30%)만이 유의한 상관관계를 보였다.(표 5)

표 5. 3점 등간척도법의 평가자내 신뢰도

	상관계수 r	
언어치료전문가	R01	0.883*
	R02	0.535
	R03	0.875
	R04	1.000***
	R05	0.423
	R06	0.764
	R07	1.000***
대학원생	R08	0.645
	R09	0.667
	R10	0.873
	R11	0.612
	R12	0.535
	R13	0.667
	R14	0.896*
	R15	0.535
	R16	1.000***
	R17	1.000***

* p<0.05, *** p<0.001

(2) 5점 등간척도법

평가자들이 5점의 등간척도법을 사용하여 말명료도를 평가한 경우에는 모든 평가자의 평가자내 신뢰도가 유의하게 나타났다. 특히 각 상관계수가 모두 0.8 이상으로 매우 높은 상관관계를 보였다. 언어치료전문가 집단의 상관계수의 평균(±표준편차)은 0.933(±0.063)이었으며, 대학원생 집단의 평균(±표준편차)은 0.964(±0.018)였다.(표 6)

표 6. 5점 등간척도법의 평가자내 신뢰도

	상관계수 r	
언어치료전문가	R01	0.884*
	R02	0.822*
	R03	0.932*
	R04	1.000***
	R05	0.942*
	R06	0.948*
	R07	1.000***
대학원생	R08	0.944*
	R09	1.000***
	R10	0.942*
	R11	0.977**
	R12	0.955*
	R13	0.976**
	R14	0.968**
	R15	0.963**
	R16	0.944*
	R17	0.970**

* p<0.05, ** p<0.01, *** p<0.001

(3) 7점 등간척도법

7점 척도를 사용했을 경우 언어치료전문가 집단에서는 모든 평가자의 평가자내 신뢰도가 유의하게 나타난 반면, 대학원생 집단에서는 10명 중 6명(60%)만 유의한 신뢰도를 가지는 것으로 나타났다. 언어치료전문가 집단의 상관계수의 평균(\pm 표준편차)은 0.955(\pm 0.076)였다.(표 7)

표 7. 7점 등간척도법의 평가자내 신뢰도

		상관계수 r
언어치료전문가	R01	0.791*
	R02	0.939*
	R03	0.962**
	R04	0.994**
	R05	1.000***
	R06	1.000***
	R07	1.000***
대학원생	R08	1.000***
	R09	0.972**
	R10	0.784
	R11	0.820
	R12	0.953*
	R13	0.956*
	R14	0.912*
	R15	0.826
	R16	0.471
	R17	0.970**

* p<0.05, ** p<0.01, *** p<0.001

나. 평가자간 신뢰도

(1) 3점 등간척도법

① 언어치료전문가 집단

언어치료전문가 집단에서는 7명의 평가자간 산출된 상관관계 계수 중 19가지 계수(90.47%)에서만 유의한 상관관계를 보였다. 평가자간 상관관계가 유의하지 않은 경우는 평가자 2와 3, 평가자 2와 4로서 각 상관계수는 $r=0.381(p>0.05)$, $r=0.351(p>0.05)$ 이었다. 유의한 상관관계를 보이지 않은 두 경우를 제외하고 가장 상관관계가 높았던 경우는 평가자 6과 7이었으며, 가장 낮은 경우는 평가자 2와 5였다. 상관계수의 분포는 $r=0.398\sim 0.770$ 이었으며, 이들의 평균(\pm 표준편차)은 $r=0.584(\pm 0.108)$ 였다.(표 8)

표 8. 3점 등간척도법에서 언어치료전문가 집단의 평가자간 신뢰도

	R02	R03	R04	R05	R06	R07
R01	0.408*	0.597**	0.643**	0.565**	0.719***	0.664***
R02		0.381	0.351	0.398*	0.562**	0.673***
R03			0.688***	0.499*	0.636**	0.614**
R04				0.415*	0.642**	0.607**
R05					0.535**	0.463*
R06						0.770***

* $p<0.05$, ** $p<0.01$, *** $p<0.001$

② 대학원생 집단

대학원생 집단의 경우 10명의 평가자에 의해 산출된 상관관계 계수 45가지에서 평가자들 간의 상관관계가 유의하게 나타난 계수는 43가지(95.56%)이다. 유의하지 않은 상관관계를 보인 경우는 평가자 8과 15($r=0.236$, $p>0.05$), 평가자 9와 15($r=0.350$, $p>0.05$)였다. 유의하지 않은 상관관계를 보인 두 경우를 제외하고, 가장 상관관계가 높았던 경우는 평가자 12와 15였으며, 가장 상관관계가 낮았던 경우는 평가자 10과 15였다. 따라서 상관계수는 $r=0.414\sim 0.900$ 까지의 분포를 보였으며 평균은 $r=0.635(\pm 0.121)$ 였다.(표 9)

표 9. 3점 등간척도법에서 대학원생 집단의 평가자간 신뢰도

	R09	R10	R11	R12	R13	R14	R15	R16	R17
R08	0.538**	0.643**	0.565**	0.422*	0.585**	0.627**	0.236	0.483*	0.745***
R09		0.740***	0.615**	0.463*	0.698***	0.728***	0.350	0.543**	0.742***
R10			0.747***	0.600**	0.643**	0.785***	0.414*	0.662***	0.724***
R11				0.685***	0.718***	0.754***	0.488*	0.805***	0.651***
R12					0.658***	0.668***	0.900***	0.674***	0.564**
R13						0.644**	0.444*	0.807***	0.639**
R14							0.496*	0.709***	0.825***
R15								0.442*	0.447*
R16									0.607**

* $p<0.05$, ** $p<0.01$, *** $p<0.001$

(2) 5점 등간척도법

① 언어치료전문가 집단

언어치료전문가 집단이 5점 등간척도법을 사용하여 말명료도를 평가했을 때, 모든 평가자들 간의 상관관계가 유의하게 나타났다. 가장 상관관계가 높은 경우는 평가자 1과 5였으며, 가장 낮은 상관관계를 보인 경우는 평가자 4와 5였다. 이들의 상관계수는 $r=0.577\sim 0.928$ 의 분포를 보였으며, 평균(\pm 표준편차)은 $r=0.741(\pm 0.109)$ 이었다.(표 10)

표 10. 5점 등간척도법에서 언어치료전문가 집단의 평가자간 신뢰도

	R02	R03	R04	R05	R06	R07
R01	0.687***	0.740***	0.642**	0.928***	0.898***	0.876***
R02		0.588**	0.721***	0.610**	0.759***	0.686***
R03			0.629**	0.675***	0.738***	0.815***
R04				0.577**	0.688***	0.705***
R05					0.911***	0.810***
R06						0.872***

** p<0.01, *** p<0.001

② 대학원생 집단

언어치료전문가 집단에서와 마찬가지로 대학원생 집단에서도 모든 평가자들 간의 상관관계가 유의하게 나타났다. 가장 상관관계가 높았던 경우는 평가자 12와 15이며 가장 상관관계가 낮았던 경우는 평가자 9와 14이다. 상관계수는 $r=0.636\sim 0.954$ 까지 분포되어 있었으며 평균(\pm 표준편차)은 $r=0.778(\pm 0.077)$ 이었다.(표 11)

표 11. 5점 등간척도법에서 대학원생 집단의 평가자간 신뢰도

	R09	R10	R11	R12	R13	R 14	R15	R16	R17
R08	0.777***	0.806***	0.815***	0.705***	0.804***	0.718***	0.648***	0.855***	0.791***
R09		0.770***	0.659***	0.754***	0.703***	0.636**	0.696***	0.781***	0.811***
R10			0.791***	0.793***	0.849***	0.718***	0.751***	0.787***	0.852***
R11				0.678***	0.856***	0.637**	0.568**	0.756***	0.742***
R12					0.857***	0.790***	0.954***	0.867***	0.795***
R13						0.851***	0.794***	0.874***	0.857***
R14							0.769***	0.874***	0.817***
R15								0.832***	0.740***
R16									0.822***

** p<0.01, *** p<0.001

(3) 7점 등간척도법

① 언어치료전문가 집단

7점 척도를 사용하여 환자의 말명료도를 평가했을 때, 경험이 많은 평가자 군에서는 모든 평가자들 간에 유의한 상관관계를 보였다. 평가자 5와 6의 상관관계가 가장 높게 나타났으며, 평가자 2와 3의 상관관계가 가장 낮았다. 상관계수는 $r=0.669\sim 0.906$ 의 분포를 보였으며, 평균(\pm 표준편차)은 $r=0.767(\pm 0.074)$ 이었다.(표 12)

표 12. 7점 등간척도법에서 언어치료전문가 집단의 평가자간 신뢰도

	R02	R03	R04	R05	R06	R07
R01	0.634**	0.669***	0.748***	0.771***	0.688***	0.769***
R02		0.669***	0.737***	0.726***	0.772***	0.755***
R03			0.818***	0.744***	0.731***	0.868***
R04				0.727***	0.787***	0.843***
R05					0.906***	0.861***
R06						0.882***

** p<0.01, *** p<0.001

② 대학원생 집단

대학원생 집단도 언어치료전문가 집단과 마찬가지로 모든 평가자들 간의 상관관계가 유의하게 나타났다. 평가자 12과 15의 상관관계가 가장 높았으며, 평가자 11과 14의 상관관계가 가장 낮았다. 이들의 상관계수는 $r=0.459\sim 0.966$ 까지 분포되어 있었으며 평균(\pm 표준편차)은 $r=0.755(\pm 0.103)$ 였다.(표 13)

표 13. 7점 등간척도법에서 대학원생 집단의 평가자간 신뢰도

	R09	R10	R11	R12	R13	R 14	R15	R16	R17
R08	0.787***	0.898***	0.720***	0.744***	0.807***	0.69***	0.729***	0.846***	0.846***
R09		0.758***	0.620**	0.552**	0.771***	0.642**	0.533**	0.742***	0.849***
R10			0.828***	0.830***	0.886***	0.754***	0.806***	0.909***	0.843***
R11				0.753***	0.815***	0.459***	0.750***	0.821***	0.709***
R12					0.738***	0.700***	0.966***	0.731***	0.652***
R13						0.740***	0.726***	0.932***	0.826***
R14							0.690***	0.703***	0.679***
R15								0.713***	0.642**
R16									0.832***

** p<0.01, *** p<0.001

3. 세 가지 등간척도법 비교

(표 14)는 세 가지 등간척도법에 대한 타당도와 신뢰도 검정 결과를 종합하여 정리해 놓은 것이다. 이 표에 의하면 언어치료전문가 집단의 경우 3점 척도는 타당도의 상관계수가 0.6 이상을 보이지 않고, 평가자내 신뢰도와 평가자간 신뢰도의 경우 유의한 수준을 만족하지 못하는 평가자가 관찰되어 좋은 평가방법이라고 할 수 없다. 5점 척도와 7점 척도는 유의한 타당도와 신뢰도의 조건을 모두 충족하므로 두 평가방법이 모두 좋은 평가방법이라고 할 수 있겠다. 반면, 대학원생 집단의 경우 3점 척도는 평가자내 신뢰도와 평가자간 신뢰도가 유의하지 않은 평가자가 관찰되며, 7점 척도 역시 평가자내 신뢰도가 유의하지 않은 평가자가 관찰되었다. 따라서 대학원생 집단에서는 5점 척도만이 좋은 평가방법으로 지적될 수 있다.

이 때 두 집단을 구분하지 않고 공통으로 사용할 수 있는 등간척도법을 제안하기 위해서는 두 집단에서 모두 유의하게 높은 타당도와 신뢰도를 만족해야 할 것이다. 따라서 5점 척도가 경험 여부에 상관없이 타당도와 신뢰도를 만족하는 유용한 평가방법이라고 할 수 있다.

표14. 세 가지 등간척도법에 대한 두 평가자 집단 간 타당도와 신뢰도 비교

	언어치료전문가			대학원생		
	타당도	평가자내 신뢰도 ¹	평가자간 신뢰도 ²	타당도	평가자내 신뢰도 ¹	평가자간 신뢰도 ²
3점 척도	-0.577	43%	90%	-0.608	30%	96%
5점 척도	-0.661	100%	100%	-0.632	100%	100%
7점 척도	-0.692	100%	100%	-0.600	60%	100%

¹ (유의한 신뢰도를 보인 평가자 수 / 각 집단의 평가자 수 × 100)

² (유의한 신뢰도를 보인 쌍의 수 / 각 집단에서 평가자간 이루어지는 쌍의 수 × 100)

IV. 고 찰

등간척도법에서 척도 간격에 대한 논의가 필요한 이유는 환자의 상태에 대해 정밀하게 평가할 수 있으면서도 높은 신뢰도를 갖는 평가방법을 찾기 위해서이다. 평가가 유의한 신뢰도를 가지려면 평가자가 평가 항목에 대한 내적 기준(internal standard)을 정립하는 것이 중요하다.³⁹ 만일 평가자가 확고한 내적 기준을 가지고 있다면 등간척도법의 등간이 세분화될수록 환자에 대한 정확한 평가가 가능할 것이다. 따라서 본 연구는 환자에 대한 정밀한 평가가 가능한 최소의 간격과 평가자의 내적 기준이 유지되는 최대의 간격이 만나는 척도 간격을 찾는 것이 목적이 된다. 여기서 환자에 대한 정밀한 평가가 가능한 지에 대한 논의는 타당도에 의해, 평가자의 내적 기준의 정립 여부는 신뢰도에 의해 평가된다. 그런데 실제 임상에서 치료 경험에 따라 평가결과에 대한 신뢰도가 달라질 수 있다. 따라서 치료 경험 여부에 따라 두 조건을 만족하는 척도 간격이 어떻게 변화되는지에 대한 논의도 포함하였다.

세 가지 등간척도법의 공인타당도와 신뢰도를 검정한 결과, 언어치료전문가 집단에서는 5점 척도와 7점 척도의 등간척도법이 모두 유의하게 높은 타당도와 신뢰도를 갖는 평가방법이었으나, 경험이 없는 대학원생 집단에서는 5점 척도만이 높은 타당도와 신뢰도를 갖는 평가방법으로 밝혀졌다. 따라서 두 집단에서 공통적으로 유용한 평가방법은 5점 척도라고 할 수 있다. 등간척도법의 점수 척도 중 5점 척도가 가장 적절한 평가방법이라는 주장은 Duffy¹에 의해 지적된 바 있다. 또한 서미경과 김향희⁴⁰도 5점 척도가 환자의 상태를 충분히 설명할 정도로 등간이 세분화되어 있으며 평가 신뢰도에도 방해를 주지 않는 좋은 평가법이라고 지적하였다.

치료경험의 차이에 상관없이 좋은 평가방법은 5점 척도이지만, 치료경

험이 쌓여 환자의 평가에 대한 내적 기준이 정립되면 7점 척도 역시 유의한 신뢰도를 유지하는 것으로 나타났다. 언어치료전문가 집단에서는 3점 척도에서 타당도와 신뢰도가 유의하지 않았으나 5점과 7점 척도에서 모두 유의한 타당도와 신뢰도를 만족하였다. 특히 5점 척도보다 7점 척도에서 더 상관관계가 높은 것으로 나타나 7점 척도가 가장 타당도와 신뢰도가 높은 평가척도로 제안되었다. 이러한 결과는 3점에서 5점, 7점 척도로 등간의 개수가 많아질수록 타당도와 신뢰도가 증가하는 선형적인 관계를 보인다고 할 수 있다. 반면, 대학원생 집단에서는 이러한 선형적 관계가 나타나지 않았다. 대학원생 집단에서는 5점 척도에서만 유의하게 높은 타당도와 신뢰도를 보이고, 3점 척도와 7점 척도에서는 각각 유의하지 않은 신뢰도가 관찰되었다. 이러한 결과는 언어치료전문가의 경우 마비말장애 환자의 말명료도 평가에 대한 경험이 많으므로 등간의 수가 증가하더라도 타당도와 신뢰도를 잃지 않으며 오히려 더 정확한 평가가 가능하다는 것을 보여준다. 그러나 대학원생의 경우 평가에 대한 경험이 부족하기 때문에 등간의 수가 너무 적거나 많으면 내적 기준이 정립되어 있지 않아 이것이 방해요인으로 작용한 것으로 보인다.

경험이 증가함에 따라 신뢰도가 증가한다는 이러한 결과는 음성평가에 대한 연구에서 제시된 바 있다. Lewis 등⁴¹과 Laczoi 등⁴²의 연구에서는 각각 언어치료전문가와 대학원생 집단의 비음평가에 대한 신뢰도를 비교하여 경험이 있는 언어치료전문가 집단의 신뢰도가 높다는 결과를 보였다. 경험과 훈련에 대한 중요성은 Eadie와 Baylor의 연구⁴³에서도 밝혀진 바 있다. 이 연구에서는 경험이 없는 평가자가 처음 평가한 결과와 이후 두 시간 가량 청지각적 훈련을 시킨 후 재평가한 결과의 신뢰도를 비교하였다. 그 결과 훈련 후의 평가자내 신뢰도와 평가자간 신뢰도가 증가하는 양상을 보였다.

등간척도법이 주관적인 평가이기 때문에 평가척도로서의 타당도가 낮다는 주장이 제기되어 왔지만 본 연구를 통해 등간척도법의 타당성을 증명하였다. 이는 청각장애 환자의 말명료도 평가에 대한 평가방법을 비교한 Samar and Metz²⁰의 연구와 같은 결과이다. 이 연구에서는 환자의 말을 평가자가 받아 적는 전사방법과 등간척도법을 비교하여 두 평가방법이 높은 상관관계를 가지는 것을 보였다. Zraick and Liss²⁷도 DME와 등간척도법에 대한 비교연구에서 두 평가방법 간에 높은 상관관계가 있다는 결과를 보였다. 따라서 본 연구는 등간척도법이 타당한 평가방법이라는 이전의 연구를 지지한다고 할 수 있다.

또한 본 연구의 가장 큰 의의는 등간척도법의 척도 점수의 종류를 비교하여 가장 타당하고 신뢰로운 척도 점수를 찾아냈다는 점이다. 그동안 이에 대한 연구가 이루어지지 않아 연구자와 언어치료전문가들이 손쉬운 방법으로 등간척도법을 사용하면서도 연구와 치료현장에서 자신이 사용하는 등간척도법의 적절성을 주장하는 근거를 충분히 마련하지 못했다. 또한 연구자들이 개념적으로 5점 척도의 유용성을 주장하였으나 그 구체적인 근거를 마련하지 못했다. 이에 본 연구는 5점 등간척도법의 타당도와 신뢰도에 대한 구체적인 근거를 마련해 주었다. 본 연구결과는 실제 치료 현장에서 언어치료전문가가 마비말장애 환자의 말명료도를 평가할 때에도 적용이 가능할 것이다. 많은 언어치료전문가들이 각각 사용하는 척도 점수가 달라 그 결과를 해석하는데 어려움이 있었으나, 본 연구에서 제안한 바와 같이 5점 등간척도법을 통일해서 사용한다면 공통된 척도 점수를 사용해 환자의 상태에 대해 논의할 수 있을 것이다. 특히 5점 등간척도법이 경험 여부에 상관없이 높은 타당도와 신뢰도를 보이므로 타 평가자의 평가내용에 대해서도 신뢰할 수 있을 것이다.

이러한 의의에도 불구하고 본 연구는 다음과 같은 제한점을 갖는다. 첫

째, 본 연구는 치료경험 여부에 따라 적절하다고 평가되는 척도 점수가 차이가 있는지에 대해 알아보았다. 그러나 치료경험 여부를 기준으로 나눈 집단의 평가자 수가 충분하지 않아 제한점으로 지적된다. 또한 치료경험을 기준으로 집단을 나누었을 때 그 집단이 좀 더 세분화 되었다면 집단간 경험의 차이가 어떤 양상을 보이는지 좀 더 자세하게 볼 수 있었을 것이다.

두 번째로 본 연구는 등간척도법의 타당도와 신뢰도에 대한 연구이지만, 그 평가 내용을 말명료도에 한정하고 있다. 말명료도를 평가할 때는 5점 척도가 가장 좋은 점수로 나타났지만, 이러한 결과가 다른 말측면에서도 동일하게 나타날 수 있는지 후속연구가 이어질 수 있다면 전반적인 측면에서의 등간척도법의 타당도와 신뢰도 연구가 이루어질 수 있을 것이다. 특히 비음이나 목소리의 높낮이 및 크기처럼 양의 값과 음의 값이 공존해 평가가 이루어지는 경우가 있다. 이 때는 주로 9점 척도, 11점 척도 등을 사용하여 0을 기준으로 '+'와 '-'로 나누어 평가하게 되는데, 이런 등간척도법에서는 타당도와 신뢰도가 어떠한 변화를 보이는지에 대해 살펴본다면 좋은 연구가 될 수 있을 것이다.

V. 결 론

본 연구는 마비말장애 환자의 말명료도 평가를 위한 척지각적 평가척도 중 등간척도법에 대한 타당도 및 신뢰도 연구이다. 본 연구에서는 3점, 5점, 7점 척도의 세 등간척도법이 치료경험의 여부에 따라 타당도와 신뢰도에 어떠한 영향을 미치는지에 대해 살펴보았다.

타당도 측면에서는 치료경험의 여부와 상관없이 세 가지 등간척도법이 모두 유의한 타당도를 가졌다. 그러나 언어치료전문가 집단이 3점 척도를 사용하여 평가한 경우에는 타당도의 상관계수가 0.6 이하로 다소 낮았다. 평가자내 신뢰도는 3점 척도의 경우 언어치료전문가 집단과 대학원생 집단에서 각각 43%와 30%의 평가자만이 유의한 신뢰도를 보였으나, 5점 척도에서는 두 집단의 평가자가 모두 유의한 신뢰도를 보였다. 7점 척도의 경우 언어치료전문가 집단의 평가자는 모두 유의한 신뢰도를 보인 반면, 대학원생 집단에서는 유의한 신뢰도를 보인 평가자가 60%에 그쳤다. 평가자간 신뢰도의 경우 언어치료전문가 집단과 대학원생 집단의 평가자 모두 3점 척도를 제외한 5점 척도와 7점 척도에서 높은 신뢰도를 보였다. 결론적으로 경험이 있는 언어치료전문가 집단에게는 5점 척도와 7점 척도가 높은 타당도와 신뢰도를 가지는 평가방법이라고 할 수 있지만, 치료 경험이 없는 대학원생 집단에게는 5점 척도만이 높은 타당도와 신뢰도를 가지는 평가방법으로 나타났다. 따라서 5점 척도가 치료경험의 여부와 상관없이 타당하고 신뢰로운 검사방법이라고 제안할 수 있다. 그리고 평가자의 치료경험이 증가하면 평가자의 신뢰도가 증가하여 7점 척도 역시 타당하고 신뢰로운 평가방법으로 제안될 수 있을 것이다. 이러한 결과는 채점자들 간의 신뢰도를 높이고 동시에 청지각적 요인의 손상 정도를 폭넓게 표현하는데

5점 척도가 가장 타당하다는 이전의 주장을 지지한다. 또한 본 연구는 경험 정도가 증가함에 따라 평가자의 신뢰도가 증가한다는 이전의 연구 결과와도 상응한다.

이상의 결과는 그동안 막연하게 5점 척도가 좋다고 주장했던 연구자들의 주장을 뒷받침해주는 구체적인 근거를 마련했다는 점에서 가장 큰 의의를 찾을 수 있다. 이에 본 연구가 앞으로 등간척도법을 사용한 연구와 치료 현장에서 도움이 될 수 있을 것이다. 연구자와 언어치료전문가들은 자신이 사용하는 등간척도 점수의 타당성에 대해 구체적인 근거를 마련할 수 있을 것이며, 제각기 사용하던 등간척도 점수를 통일하는데도 도움이 될 것이다. 그러나 본 연구는 등간척도법에 대한 연구를 말명료도 측면으로 한정하였으므로 다른 말 측면에서도 같은 결과가 나타나는지에 대한 후속 연구가 필요하다.

참 고 문 헌

- 1) Duffy JR. Motor speech disorders: substrates, differential diagnosis, and management. Mayo foundation for medication and research; 1995.
- 2) Darley FL, Aronson AE, Brown JR. Differential diagnostic patterns of dysarthria. J Speech Hear Res 1969a;12:246-69.
- 3) Darley FL, Aronson AE, Brown JR. Clusters of deviant speech dimensions in the dysarthrias. J Speech Hear Res 1969b;12:462-9.
- 4) Cahill LM, Murdoch BK, Theodoros DG. Perceptual analysis of speech following traumatic brain injury in childhood. Brain Inj 2002;16:415-36.
- 5) Weismer G, Jeong JY, Laures JS, Kent RD, Kent JF. Acoustic and intelligibility characteristics of sentence production in neurogenic speech disorders. Folia Phoniatr Logop 2001;53:1-18.
- 6) Yorkston KM, Beukelman DB, Bell KR. Clinical management of dysarthria speakers. Boston: College-Hill Press; 1988.
- 7) Platt LJ, Andrews G, Young M, Neilson PD. The measurement of speech impairment of adults with cerebral palsy. Folia Phoniatr Logop 1978;30:50-8.
- 8) 김수진. 말운동장애의 평가. 한국언어청각임상협회 여름연수회 2001;71-9.
- 9) Kent RD. Intelligibility in speech disorders. Philadelphia, PA: John Benjamins; 1992.
- 10) Weston A, Shiriberg L. Contextual and linguistic correlates of

- intelligibility in children with developmental phonological disorders. *J Speech Hear Res* 1992;35:1316-32.
- 11) 윤미선, 이승환. 정상 및 기능적 조음장애 아동의 자음정확도와 명료도 검사방법의 비교. *언어청각장애연구* 1998;3:50-65.
 - 12) Hodson B, Paden E. Phonological processes which characterize unintelligible and intelligible speech in early childhood. *J Speech Hear Res* 1981;46:369-73.
 - 13) Osberger MJ. Scaling procedures for the measurement of speech intelligibility. In: Kent RD, editor. *Intelligibility in speech disorders*. Philadelphia: John Benjamins; 1992. p.233-63.
 - 14) Bernthal J, Bankson N, *Articulation and phonological disorders*. Englewood Cliff, NJ: Prentice-Hill; 1998.
 - 15) Schiavetti N. Scaling procedures for the measurement of speech intelligibility. In: Kent RD, editor. *Intelligibility in speech disorders*. Philadelphia: John Benjamins; 1992. p.11-34.
 - 16) 김향희, 이미숙, 김선우, 이원용. 파킨슨증으로 인한 마비말장애의 청지각적 평가에 대한 신뢰도 연구. *음성과학* 2004;11:129-41.
 - 17) Chan KM, Yiu EM. The effect of anchors and training on the reliability of perceptual voice evaluation. *J Speech Lang Hear Res* 2002;45:111-26.
 - 18) Kempler D, Lancker DV. Effect of speech task on intelligibility in dysarthria: a case study of parkinson's disease. *Brain Lang* 2002;80:449-64.
 - 19) Whitehill TL, Lee SY, Chun JC. Direct magnitude estimation and interval scaling of hypernasality. *J Speech Lang Hear Res* 2002;45:80-8.

- 20) Samar VJ, Metz DE. Criterion validity of speech intelligibility rating-scale procedures for the hearing-impaired population. *J Speech Hear Res* 1998;31:307-16.
- 21) Kent RD. Speech intelligibility and communicative competence in children. In: Kariser AP, Gray DB, editors. *Enhancing children's communication: research foundation for intervention*. Baltimore; 1993.
- 22) Hustad KC. Estimating the intelligibility of speakers with dysarthria. *Folia Phoniatr Logop* 2006;58:217-28.
- 23) Weismer G, Laures JS. Direct magnitude estimates of speech intelligibility in dysarthria: effects of a chosen standard. *J Speech Lang Hear Res* 2002;45:421-33.
- 24) 윤미선, 이윤경, 심현섭. 청각장애아동의 말명료도에 영향을 미치는 화자 요인. *언어청각장애연구* 2000;5:144-58.
- 25) 김수진. 일음절 낱말대조 명료도 평가방법을 이용한 마비말장애의 분절적 특성 연구: 경직형과 이완형의 비교. 이화여자대학교 대학원. 언어병리협동과정 박사학위청구논문; 2000.
- 26) 김향희, 이미숙, 김선우, 최성희, 이원용. 파킨슨증으로 인한 마비말장애에 대한 청지각적 평가척도. *음성과학* 2004;11:39-49.
- 27) Zraick RI, Liss JM. A comparison of equal-appearing interval scaling and direct magnitude estimation of nasal voice quality. *J Speech Lang Hear Res* 2000;43:979-88.
- 28) Hoodin RB, Gilbert HR. Parkinsonian dysarthria: an aerodynamic and perceptual description of velopharyngeal closure for speech. *Folia Phoniatr Logop* 1989;41:249-58.
- 29) Ludlow CL, Bassich CJ. Relationships between perceptual ratings

- and acoustic measures of hypokinetic speech. In McNeil MR, Rosencek JC, Aronson AE, editors. The dysarthria: physiology, acoustics, perception, management. San Diego: College-Hill Press; 1984.
- 30) Karnell MP, Folkins JW, Morris HL. Relationships between the perception of nasalization and speech movements in speakers with cleft palate. J Speech Hear Res 1985;28:63-72.
- 31) Workinger MS, Kent RD. Perceptual analysis of the dysarthria in children with athetoid and spastic cerebral palsy. In: Moore CA, Yorkston KM, and eukelman DR, editors. Dysarthria and apraxia of speech: perspectives on management; 1991. p.109-26.
- 32) Hathaway RR. A comparison of transverse and vertical pharyngeal flaps using electromyography and judgements of nasality. Cleft Palate J 1980;17:305-8.
- 33) Stevens SS. Psychophysics. New York: Weley; 1975.
- 34) 성태제. 타당도와 신뢰도. 학지사; 1995.
- 35) 김현승. 마비말장애 환자의 파열음 산출의 음향학적 분석. 연세대학교 언어병리학협동과정 석사학위 논문; 2006.
- 36) 김향희, 나덕렬. 파다라이스 · 한국판-웨스턴 실어증검사. 서울: 재단법인 파라다이스복지재단; 2001.
- 37) 김향희. 신경 말 · 언어장애. 의사소통장애개론. 서울: 하나의학사; 2001. p.127-50.
- 38) Sheard C, Adams RD, Davis PJ. Reliability and agreement of ratings of ataxic dysarthric speech samples with varying intelligibility. J Speech Hear Res 1991;34:285-93.

- 39) Kreiman J, Gerratt BR, Kempster GB, Erman A, Berke GS. Perceptual evaluation of voice quality: review, tutorial, and a framework for research. *J Speech Hear Res* 1993;36:21-40.
- 40) 서미경, 김향희. 마비말장애 연구문헌에서 살펴본 말평가의 청지각적 요소. *음성과학* 2006;13:197-206.
- 41) Lewis KE, Watterson TL, Houghton SM. The influence of listener experience and academic training on ratings of nasality. *J Commun Disord* 2003;36:49-58
- 42) Laczi E, Sussman JE, Stathopoulos ET, Huber J. Perceptual evaluation of hypernasality compared to HONC Measures: the role of experience. *Cleft Palate Craniofac J* 2005;42:202-11.
- 43) Eadie TL, Baylor CR. The effect of Perceptual training on inexperienced listeners' judgments of dysphonic voice. *J Voice* 2006;20:527-44.

부록-1. 일음절 낱말대조 명료도 검사의 문항

음절위치	대조양식	전체대조군	문항	대조내용
초성	조음위치	73-74	칼: 팔	연구개폐쇄: 양순폐쇄
		47-48	손: 혼	치조마찰: 후두마찰
		61-62	담: 감	경구개폐쇄: 연구개폐쇄
		65-66	말: 날	양순비음: 치조비음
		71-72	땅: 빵	치조폐쇄: 양순폐쇄
	조음방법	23-24	산: 난	치조마찰: 치조비음
		43-44	살: 달	치조마찰: 치조폐쇄
		55-56	학: 악	후두마찰: 개방음
		63-64	망: 방	양순비음: 양순폐쇄
		67-68	딸: 쌀	치조폐쇄: 치조마찰
	상관속 (발성방법)	87-89	뿔: 불: 풀	양순폐쇄(연): 양순폐쇄(경)
		17-18	살: 쌀	치조마찰(연): 치조마찰(경)
		96-98	달: 딸: 탈	치조폐쇄(경): 치조폐쇄(격)
		93-95	침: 짐: 짐	경구개과찰(경): 경구개과찰(격)
		90-92	공: 룡: 풍	연구개폐쇄(연): 연구개폐쇄(격)
	위치+방법	13-14	산: 잔	치조마찰: 경구개과찰
		53-54	잠: 남	경구개과찰: 치조비음
		69-70	쌍: 짱	치조마찰: rudrnrovkckf
		81-82	발: 날	양순폐쇄: 치조비음
		31-32	힘: 짐	후두마찰: 연구개폐쇄
1-2		공: 중	연구개폐쇄: 경구개과찰	
41-2		김: 짐	연구개폐쇄: 경구개과찰	
중성	고저	29-30	힘: 험	고설전설: 저설전설
		9-10	쌈: 심	고설전설: 저설전설
		5-6	국: 각	고설후설: 저설후설
		49-50	달: 들	저설후설: 고설후설

전 후	3-4	뱀: 밤	저설 전설: 저설 후설	
	25-26	상: 생	저설 전설: 저설 후설	
	35-36	징: 중	고설 후설: 고설 전설	
	7-8	물: 밀	고설 후설: 고설 전설	
고저+전 후	51-52	갓: 갓	고설 전설: 저설 후설	
	37-38	짐: 잠	고설 전설: 저설 후설	
	45-46	축: 책	저설 전설: 고설 후설	
	79-80	댁: 독	저설 전설: 고설 후설	
중 성	조음 위치	33-34	밖: 밥	연구개 폐쇄: 양순 폐쇄
		75-76	답: 닥	연구개 폐쇄: 양순 폐쇄
		77-78	당: 담	양순비음: 연구개비음
		57-58	밤: 방	양순비음: 연구개비음
	조음 방법	11-12	밤: 밤	연구개 폐쇄: 양순비음
		15-16	밭: 발	치조 폐쇄: 치조유음
		19-20	밭: 반	치조유음: 치조비음
		39-40	밭: 반	치조 폐쇄: 치조비음
		59-60	밖: 방	연구개 폐쇄: 연구개비음
	위 치+방법	83-84	방: 밥	연구개비음: 양순 폐쇄
		27-28	당: 달	연구개비음: 치조유음
		85-86	밖: 발	연구개 폐쇄: 치조유음
21-22		밤: 발	양순비음: 치조유음	

출처: 김수진(2001) 일음절 낱말대조 평가

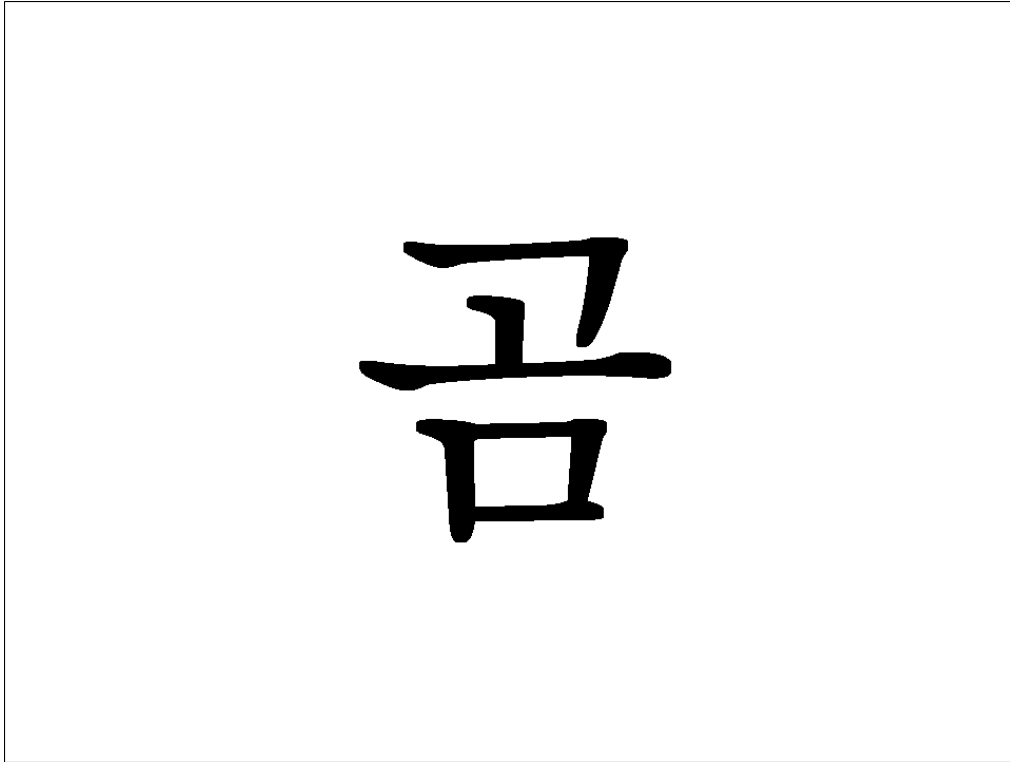
부록-2. 일음절 낱말대조 명료도 검사의 문항 평가지

예1	곰	봄	숨	몸	좁	예2	곰	봄	숨	몸	좁
예3	검	컴	검	예4	검	겁	검	예5	검	컴	검
1	공	총	동	콩	종	2	공	총	동	콩	종
3	뱀	밤	봄	븀	범	4	뱀	밤	븀	븀	범
5	객	국	곡	각	극	6	객	국	곡	각	극
7	실	살	술	술	설	8	실	살	술	술	설
9	심	삼	쌈	숨	숨	10	심	삼	쌈	숨	숨
11	밤	반	밭	밥	방	12	밤	반	밭	밥	방
13	산	단	난	잔	한	14	산	단	난	잔	한
15	밭	반	밭	방	밤	16	밭	반	밭	방	밤
17	탈	살	달	쌀	딸	18	탈	살	달	쌀	딸
19	밭	밭	반	방	밥	20	밭	밭	반	방	밥
21	밭	반	밭	방	밤	22	밭	반	밭	방	밤
23	산	단	잔	난	간	24	산	단	잔	난	간
25	성	상	생	승	송	26	성	상	생	승	송
27	당	단	달	담	닷	28	당	단	달	담	닷
29	흙	힘	힘	힘	힘	30	흙	힘	힘	힘	힘
31	검	힘	점	섬	덤	32	검	힘	점	섬	덤
33	반	밖	밥	방	밭	34	반	밖	밥	방	밭
35	장	징	중	중	정	36	장	징	중	중	정
37	잠	점	짐	젼	좁	38	잠	점	짐	젼	좁
39	반	밭	방	밭	밤	40	반	밭	방	밭	밤
41	김	님	팀	짐	침	42	김	님	팀	짐	침
43	탈	살	달	딸	쌀	44	탈	살	달	딸	쌀
45	책	척	측	축	축	46	책	척	측	축	축
47	돈	손	논	혼	본	48	돈	손	논	혼	본
49	달	덜	들	돌	돌	50	달	덜	들	돌	돌
51	곳	깃	갓	것	굿	52	곳	깃	갓	것	굿
53	담	삼	남	잠	감	54	담	삼	남	잠	감
55	악	각	학	닭	삿	56	악	각	학	닭	삿
57	밭	반	밤	방	밖	58	밭	반	밤	방	밖
59	밭	반	밤	방	밖	60	밭	반	밤	방	밖

61	담	밤	감	잠	삼	62	담	밤	감	잠	삼
63	상	강	방	망	당	64	상	강	방	망	당
65	달	살	날	말	발	66	달	살	날	말	발
67	쌀	칼	딸	팔	탈	68	쌀	칼	딸	팔	탈
69	각	빽	째	딱	짹	70	각	빽	째	딱	짹
71	깡	뺑	짱	땅	쌍	72	깡	뺑	짱	땅	쌍
73	탈	팔	알	칼	살	74	탈	팔	알	칼	살
75	달	당	닭	답	담	76	달	당	닭	답	담
77	달	당	닭	답	담	78	달	당	닭	답	담
79	닭	덕	득	득	दै	80	닭	덕	득	득	दै
81	말	날	달	발	살	82	말	날	달	발	살
83	발	밤	박	방	밤	84	발	밤	박	방	밤
85	발	밤	밖	방	밤	86	발	밤	밖	방	밤
87	불	풀	뿔			88	불	풀	뿔		
90	공	콩	꿍			91	공	콩	꿍		
93	짐	침	짚			94	짐	침	짚		
96	달	탈	딸			97	달	탈	딸		
						98	달	탈	딸		

출처: 김수진(2001) 일음절 낱말대조 평가

부록-3. 말명료도 피검자에게 제시되는 카드 예



부록-4. 말명료도 피검자에게 제시되는 문단

우리 나라의 가을은 참으로 아름답다. 무엇보다도 산에 오를 땐 더욱 더 그 빼어난 아름다움이 느껴진다. 쓰다듬어진 듯한 완만함과 깎아놓은 듯한 뾰족함이 어우러진 산등성이를 따라 오르다보면 절로 감탄을 금할 수가 없게 된다. 붉은 색, 푸른 색, 노란 색 등의 여러 가지 색깔들이 어우러져 타는 듯한 감동을 주며 나아가 신비롭기까지 하다. 숲 속에 누워서 하늘을 바라보라. 쌍쌍이 짝지어 있는 듯한 흰 구름, 높고 파란 하늘을 쳐다보고 있노라면, 과연 옛부터 가을을 천고마비의 계절이라 일컫는 이유를 알게 될 것만 같다. 가을에는 또한 오곡백과 등 먹거리가 풍성하기 때문에 결실의 계절이라고도 한다. 햅쌀, 밤, 호두뿐만 아니라, 대추, 여러 가지 떡, 크고 작은 과일들을 맛볼 수 있는데, 가을의 대표적인 명절인 추석에 우리는 이것들을 쌓아놓고 조상님들께 차례를 지내기도 한다. 또한, 가을은 독서의 계절이라 하여 책을 읽으며 시시때때로 명상에 잠기기도 하는데, 독서는 우리에게 마음을 살찌우고 아름답게 하는 힘을 주기 때문이다.

출처 : 김향희(2001), '가을' 문단

부록-5. 말명료도 평가자에게 제시되는 문단

쓰다듬어진 듯한 완만함과 깎아놓은 듯한 뾰족함이 어우러진 산등성이를 따라 오르다보면 절로 감탄을 금할 수가 없게 된다. 붉은 색, 푸른 색, 노란 색 등의 여러 가지 색깔들이 어우러져 타는 듯한 감동을 주며 나아가 신비롭기까지 하다. 숲 속에 누워서 하늘을 바라보라. 쌍쌍이 짝지어 있는 듯한 흰 구름, 높고 파란 하늘을 쳐다보고 있노라면, 과연 옛부터 가을을 천고마비의 계절이라 일컫는 이유를 알게 될 것만 같다.

출처 : 김향희(2001), '가을' 문단 중 일부

Abstract

*Validity and Reliability of Equal-Appearing Interval Scales
for Speech Intelligibility Test of Dysarthria*

HyeJin Park

Graduate Program in Speech and Language Pathology, Yonsei University

(Directed by Professor HyangHee Kim)

Speech intelligibility is a measure of the degree to which a person's speech can be understood by a listener. Speech intelligibility is a measure of functional limitation, providing global information regarding how well the speech subsystems work together to produce speech. The most common method of assessing dysarthric speech intelligibility is the use of Equal-Appearing Interval (EAI) scales. EAI scales require listeners to assign a number between 1 and n to a speech sample, where n is the number of points on the scale. In previous studies, speech language pathologists (SLPs) used various point scales with no rationale given to validate the choice. The purpose of this study was to compare different rating scales to investigate the most valid and reliable scale among 3-, 5-, and 7-point EAI.

Seventeen raters participated in this study. Seven raters were licensed SLPs with experience in assessment and treatment of dysarthria. The other 10 were graduate program in speech and language

pathology with no experience in speech therapy. Twenty-five dysarthric subjects showed various levels of dysarthric severity. Each subject was evaluated with the Phonemic Contrast Test (PCT) which utilizes a closed set of response choice method. Each subject was to read a Korean standard passage, 'Autumn' and the speech intelligibility of each subject was compared with raw scores of PCT performances to investigate validity using Pearson's correlations. The ratings were repeated after one week for assessing intra-judge reliability. Pearson's correlations were calculated for intra- and inter-judge reliabilities.

The results were as follows:

1. In terms of concurrent validity, all four rating scale methods were significantly correlated with PCT.

2. In terms of intra-judge reliability, only 43% of SLP raters and 30% of the graduate student raters showed significant reliability in the 3-point scale. Using the 5-point scale, all raters of two groups showed significant reliability. In the 7-point scale, all of the SLPs and only 60% of the graduate students were significantly reliable.

3. In terms of inter-judge reliability, both groups had significant reliability using the 5- and 7-point scales. However, two pairs of each group did not show significant reliability using the 3-point scale.

In conclusion, both 5- and 7-point scales in the SLP group, and only the 5-point scale in the graduate student group were valid and reliable. Therefore, the 5-point scale was the most valid and reliable scale regardless of experience. Furthermore, as raters have more

experience, the 7-point scale can be considered as a valid and reliable method as well. These results supported other research that considered the 5-point scale of EAI as a widespread and reliable method. Further studies are needed to apply these results to other speech dimensions such as hypernasality and monopitch.

Key words: dysarthria, speech intelligibility, perceptual assessment, Equal-Appearing Interval(EAI) scales.