

The perceptual judgment of sound prolongation: Equal-appearing interval and direct magnitude estimation*

Jin Park** · Hwajung Cha · Sejin Bae

Department of Speech Language Rehabilitation, Catholic Kwandong University, Gangneung, Korea

Abstract

This study aimed to propose an appropriate evaluation method for the perceived level of speech disfluency based on sound prolongation (i.e., increased duration of segments). To this end, 34 Korean-speaking adults (9 males, 25 females, average age: 32.9 yrs.) participated as raters in this study. The participants listened to sentences containing a total of 25 stimuli with the Korean voiceless fricative /s/ extended by 80-ms increments up to 2,000 ms (i.e., 285 ms, 365 ms, ..., 2,125 ms, 2,205 ms), and evaluated them using an equal-appearing interval scale (EAI, 1–7 points, where 1 represents “normal” and 7 represents “severe”). Subsequently, based on the interval-scale results, the sentence stimuli with the prolonged voiceless fricative corresponding to the mild-to-moderate level (rated as 4 points) were selected as the reference modulus for direct magnitude estimation (DME). After scatter plots were created for the two evaluation results, the relationship between the two measured mean values was analyzed using a curve estimation method for the observed data with the highest R²-value to determine whether a linear or curvilinear approximation fit the data better. A curvilinear relationship between the two evaluation results was indicated, suggesting that DME is a more appropriate evaluation method than the EAI scale for assessing the perceived level of disfluency based on sound prolongation.

Keywords: sound prolongation, equal appearing interval scale, direct magnitude estimation, auditory perceptual evaluation, stuttering (disfluency)

1. 서론

청지각적 평가는 언어재활 임상에서 말·언어 샘플을 듣고 장애의 유무, 유형, 심각 정도를 파악할 수 있는 일반적인 측정 방식으로 통용된다(Ha, 2009). 하지만 청지각적 평가는

수행이 용이한 반면, 기본적으로 평가자의 지각에 의존하기에 주관적이며 상황에 따라 판단이 달라질 수 있는 가변적인 특성을 보일 수 있다. 따라서 정확한 진단을 위해서는 장애에 대한 숙련된 임상가의 청지각적 판단 능력이 요구된다 고 하겠다. 또한 장애의 심각 정도를 타당하고 신뢰롭게 평

* This research was supported by “Regional Innovation Strategy (RIS)” through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (MOE) in 2023 (2022RIS-005).

** gatorade70@cku.ac.kr, Corresponding author

Received 15 August 2023; Revised 11 September 2023; Accepted 11 September 2023

© Copyright 2023 Korean Society of Speech Sciences. This is an Open-Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

가하기 위해서는 적절한 평가 방법을 선택해 사용해야 할 것이다. 이러한 맥락에서 현재까지 말 명료도, 유창성 정도, 음성의 질, 공명 문제 영역 등을 포함해 다양한 측면에서의 타당한 청지각적 평가 방법을 제시하고자 하는 임상적 연구들이 진행되었다(Eadie & Doyle, 2002; Ha, 2009; Kreiman & Gerratt, 1998; Lee et al., 2022; Schiavetti et al., 1994; Weismer & Laures, 2002; Whitehill et al., 2002; Zraick & Liss, 2000).

임상에서 사용되는 보편적인 청지각적 평가 방법 중 하나인 등간척도법(equal-appearing interval scale, EAI)은 특정한 말 관련 특징의 정도를 하나의 연속선상에 일정한 간격으로 배치된 숫자 가운데 선택해 평가하는 방식이다(Schiavetti, 1992). 주로 활용되는 숫자 간격으로는 5점, 7점, 9점 등의 홀수 척도가 사용되며, 대체로 가장 낮은 수인 0이나 1을 정상(normal) 수준으로 하고, 장애가 가장 심한 정도를 가장 높은 숫자로 표시한다(Duffy, 1995). 말 명료도와 음성의 질, 유창성 정도, 공명 문제 등을 포함한 말 장애 영역에서 다양한 척도 점수를 사용하고 있다. 예를 들어, 합성된 모음(synthesized vowel)의 비성 정도를 측정해 Zraick & Liss(2000)와 파킨슨병 환자의 과대비성 정도를 측정해 Hoodin & Gilbert(1989)에서는 5점 척도를 사용하였다. 한편, 마비말장애의 여러 하위 유형을 구별하기 위한 말과 음성 관련 변수 총 38개(예, 말 속도, 음도, 강도, 음성의 질 등)의 심각 정도를 평가한 Darley et al.(1969), 파킨슨병 환자의 약한 음성 정도(hypophonia)를 측정해 Ludlow & Bassich(1984), 그리고 구개열 환자의 과대비성 정도를 평가한 Karnell et al.(1985)은 모두 7점 척도를 사용하였다. 지연청각피드백(delayed auditory feedback, DAF) 상황에서의 말더듬 성인 발화의 자연스러움 정도를 측정해 Martin et al.(1984)에서는 9점 척도로 평가하였다. EAI는 용이성으로 인해 임상에서 보편적으로 사용되는 평가 방법임에도 타당도와 관련해서 몇몇 제한점들이 제기되었다(Schiavetti et al., 1981; Schiavetti, 1992; Schiavetti et al., 1983; Stevens, 1974). 평가 결과와 관련해 특정 정도에 체계적으로 편중되어(systematic bias) 나타나는 경향성을 보이며, 일정한 구간 내의 숫자를 사용하기에 세밀한 평가 수행에 어려움이 있을 수 있다(Stevens, 1974). 또한 이로 인한 통계 분석의 제한도 있을 수 있다(Price et al., 2012).

임상에서 사용되는 또 다른 청지각적 평가 척도로서 직접 크기평정법(direct magnitude estimation, DME)이 있다. 이는 기본적으로 평가 기준치(modulus)를 통한 평가로, 평가자는 평가 기준치에 해당하는 음성 샘플을 먼저 듣고 이후 평가하는 음성 샘플에 점수를 부여하는 방식이다. 이때 평가 기준치에 특정 점수(일반적으로 100)가 부여되고 이와 비교해 평가하는 음성 샘플의 상대적 점수를 매기면 된다. 하지만 평가 기준치를 사용하지 않을 수도 있는데, 이때는 평가자가 첫 번째로 들은 음성 샘플에 임의의 점수를 부여하고 이후 음성 샘플에 대해서 그 첫 번째 샘플과 비교해 점수를 부여한다. DME는 EAI에 비해 사용이 용이하지 않다는 단점이 있으나, EAI에서 나타나는 체계적인 점수 편중화 현상을

방지할 수 있으며, 특정 평가 영역의 상대적인 정도를 좀 더 세밀하게 평가할 수 있다는 장점이 있다(Stevens, 1974). Eadie & Doyle(2005a)에서 음성장애 환자에게 대한 음성장애의 심각 정도와 음성의 듣기 좋은 정도(pleasantness)를, Weismer & Laures(2002)에서는 마비말장애 환자를 대상으로 말명료도를, Martin(1965)은 말더듬 성인을 대상으로 말더듬의 심각 정도를 모두 DME로 평가하였다.

이러한 EAI와 DME의 상대적인 타당도와 관련해 말장애의 여러 영역에서 두 평가법의 비교 분석을 통해 어떤 평가법이 적절한지에 대한 연구들이 진행되었다(Eadie & Doyle, 2002; Ha, 2009; Kreiman & Gerratt, 1998; Lee et al., 2022; Schiavetti et al., 1994; Weismer & Laures, 2002; Whitehill et al., 2002; Zraick & Liss, 2000). 두 평가법에 대한 비교 연구는 기본적으로 양적(prothetic) 또는 질적(metathetic) 연속성(continuum)이라는 심리물리학(psychophysics) 개념을 바탕으로 진행되었다(Stevens, 1974). 먼저, 양적 연속성은 강도(loudness)의 경우처럼 기존 자극에 자극이 양적으로 더해지는 부가적인(additive) 속성을 띠는 반면, 질적 연속성은 음도(pitch)의 예처럼 일정한 위치에서 다른 위치로의 질적인 변화로(예, ‘미’ 음계에서 ‘파’ 음계로 변환) 치환적인(substitutive) 속성을 나타낸다. 특정 평가 영역이 양적 연속성 또는 질적 연속성을 보이는지를 판단하기 위해서는 동일한 평가 영역을 각각 EAI와 DME의 평가 절차에 따라 평가한 후, 두 평가의 결과치가 선형적인(linear) 또는 비선형적인(curvilinear) 관계를 보이는지를 살펴봄으로써 가능하다. 만약 두 평가 결과치의 관계가 비선형으로 나타난다면 EAI보다 DME가 타당한 척도라는 것을, 반면에 두 평가 결과치의 관계가 선형적으로 나타난다면 두 평가법 모두 적절한 것으로 판명될 수 있다. 이런 방식으로 결국 청각장애인의 말명료도(Schiavetti et al., 1981), 말더듬의 심각 정도(Schiavetti et al., 1983), 연장 모음의 거친 정도(Toner & Emanuel, 1989), 과대비성 정도(Ha, 2009; Whitehill et al., 2002; Zraick & Liss, 2000), 음성장애의 심각 정도(Lee et al., 2022)에서는 DME가 타당한 평가법으로 제안되었다. 반면에, 아동의 조음 결합 정도(Prather, 1960)와 정상 화자의 기식 정도(Sewall et al., 1999)의 경우에는 두 평가법 모두 타당한 것으로 나타났다. 하지만 말의 자연스러움의 경우, Metz et al.(1990)에서는 질적인 속성을 보인 반면, Southwood(1996)에서는 양적인 속성으로 나타나 DME가 타당하다는 상반된 결과가 나타나기도 하였다.

본 연구에서는 기본적으로 연장음의 길이에 따른 비유창성 정도 인식과 관련해 EAI와 DME의 평가 결과치 간의 관계를 분석해 어떤 평가법이 타당한지를 알아보고자 하였다. 현재까지 주로 연장음에 대한 연구는 특정 분절음이 어느 정도 길어질 때 비유창하게 지각하는지 또는 점차적으로 길어질수록 비유창하게 지각하는 양상(즉, 단속적 vs. 연속적)은 어떤지에 대한 주제로 진행되었다. 예를 들어, Jones et al.(2005)에서는 분절음(영어 /z/, /a/)이 원래 길이의 38% 이

상 길어질 때 비유창하게 지각한다고 제안하였다. 영어 화자를 대상으로 한 Kawai et al.(2007)과 한국어 화자를 대상으로 한 Park & Chung(2023)에서는 분절음(즉, 영어 /f/, 한국어 /s/)이 점차적으로 길어질 때 단속적이 아닌 연속적인 지각 양상을 보인다고 주장하였다. 특히, 두 연구에서 연장음의 길이에 따른 비유창성 지각 양상에 대한 평가를 1-100점 척도(100점에 가까울수록 비유창하게 지각하는 정도가 높음)를 이용하였다. 하지만 연장음의 길이에 따른 비유창성 정도에 대한 적절한 평가를 위해서는 먼저 이러한 지각 특성이 양적 또는 질적 연속성을 보이는지 알아볼 필요가 있다. 또한 Schiavetti et al.(1983)에서는 말더듬의 심각 정도를 평가함에 있어 EAI보다는 DME가 타당한 평가법이라 제안하였다. 하지만 말더듬의 핵심행동 중 하나인 연장은 반복과 구별되는 ‘비운율적 발성’(disrhythmic phonation)의 한 형태로, 일반적으로 말더듬 성인에게 있어 막힘이나 반복보다 비교적 적은 빈도수로 관찰되는 비유창성 형태이다 (Jeon & Jeon, 2015). 따라서 막힘이나 반복과 같은 비유창성이 주로 나타나는 음성샘플을 이용한 Schiavetti et al.(1983)과는 다른 결과가 나타날 가능성도 배제할 수 없다. 이러한 동기로 본 연구에서는 연장음의 길이에 따른 비유창성 지각 정도에 있어 EAI와 DME 중 어떤 평가 방법이 적절한지를 알아보고자 한다.

2. 연구방법

2.1. 연구대상

본 연구의 청지각적 평가자로서 한국어를 모국어로 하는 만 19세 이상의 성인 대학생 남녀 34명(남: 9명, 여: 25명, 평균연령: 32.9세)이 참여하였다. 대상자 선정기준으로는 정상 시력과 청력을 가지고 있으며 실험수행에 부정적 영향을 미칠 수 있는 심리적, 정서적, 신경학적 병력이 없는 대상자로 하였다. 또한 언어재활(치료) 비전공자로 이전에 말 또는 언어 관련 청지각적 평가에 참여한 경험이나 임상 경험이 없는 대상자로 선정하였다. 마지막으로 말더듬에 대한 전문 지식이나 말을 더듬는 사람에 대한 개인적 경험(1개월 이상)이 없어야 한다. 청력검사는 순음청력검사를 실시하였으며 모두 정상역치(250 Hz와 4,000 Hz 구간에서 25 dB 이하)의 청력을 보였다. 본 실험에 앞서 미리 준비한 질문지를 통해 연령, 성별, 교육정도와 같은 기본정보와 함께 말더듬 관련 정보[예, 본인의 말더듬 유무, 말더듬 가족력 유무, 말더듬(인)에 대한 경험 유무]를 조사하였다. 조사 결과, 본인이 말을 더듬거나 말더듬 가족력을 가지고 있는 대상자는 존재하지 않았다.

2.2. 음성샘플

본 연구의 청지각적 평가에 쓰이는 음성샘플은 ‘가을문단’(Kim, 1996)의 “우리나라의 가을은 참으로 아름답다. 산에 오를 땐 더욱더 그 빼어난 아름다움이 느껴진다”에서

‘산’(/san/)의 초성인 평마찰음 /s/를 원길이의 40%에 해당되는 80 ms씩 연장해 총 25개 음성샘플을 생성하였다(즉, 285 ms, 365 ms, 2,125 ms, 2,205 ms). 서론에서 기술한 것처럼, 분절음의 원래 길이의 38% 이상 연장했을 때 비유창하게 지각한다는 Jones et al.(2005)의 결과에 따라 본 연구에서도 평마찰음 /s/의 원길이의 40%에 해당되는 80 ms를 연장해 음성샘플을 생성하였다. 변조하지 않은 원래 평마찰음 /s/의 지속시간은 205 ms이며, 마찰구간은 95 ms이고 이후의 기식구간은 110 ms이다. 마찰구간과 기식구간 모두 정상 범위에서 관찰되었다(Lee, 2001). 변조방법은 먼저 원래 평마찰음의 마찰구간이 안정적으로 발생하는 시점을 기준으로 80 ms를 복사하여 마찰구간의 정중앙에 다시 삽입하는 방식으로 285 ms의 평마찰음을 생성하였다. 또한 생성된 285 ms의 평마찰음에서 안정적인 마찰구간을 기준으로 80 ms를 복사하여 마찰구간 정중앙에 다시 삽입하여 365 ms의 평마찰음을 생성하였다. 동일한 방식으로 평마찰음 /s/의 길이가 445-2,205 ms까지의 총 25개의 음성샘플을 생성하였다. 그림 1은 평마찰음 /s/의 길이가 각각 285 ms와 2,205 ms의 청각자극을 보여주고 있다.

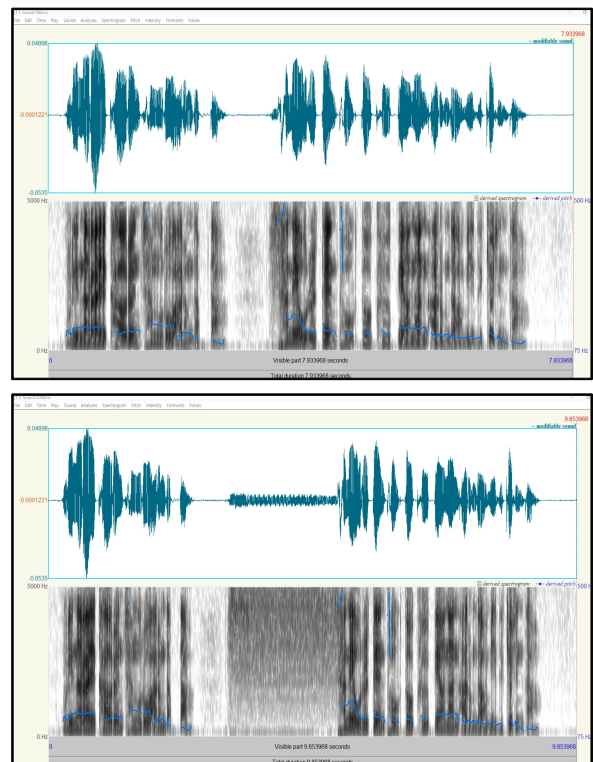


그림 1. 실험자극 예시[위: 80 ms 연장한 /s/(전체 길이: 285 ms), 아래: 2,000 ms 연장한 /s/(전체 길이: 2,205 ms)]
Figure 1. Examples of sentence stimuli [top: a stimulus elongated by 80 ms (total duration of /s/ in ‘san’=285 ms), bottom: a stimulus elongated by 2,000 ms (total duration of /s/ in ‘san’=2,205 ms)]

2.3. 연구절차

실험참여자는 동일한 음성샘플을 일주일 간격으로 각각

등간척도법(EAI)과 직접크기평정법(DME)을 이용해 두 번 평가하였다. 먼저 EAI 평가에서는 제시된 각각의 연장을 듣고 실험참여자가 연장의 길이에 따른 비유창성 정도에 대해 7점 척도[1=정상(normal), 2=미비(minimal), 3=경도(mild), 4=경중도(mild to moderate), 5=중도(moderate), 6=중심도(moderate to severe), 7=심도(severe)]로 평가하였다. 이를 통해 연장의 길이에 따른 비유창성 인식 정도와 관련해 ‘경중도’(4점)에 해당하는 음성 샘플을 정하였다. 결국 ‘경중도’(4점)의 최빈값(mode)인 1,005 ms을 선정하였다. DME 평가에서는 이를 기준 평가치(modulus)로 설정하여 이후 제시하는 각각의 연장음 샘플에 대해 비유창성 정도를 평가하도록 하였다. 즉, 기준 평가치에 해당하는 연장의 길이에 따른 비유창성 정도가 100점이라면 이후 제시되는 연장음 자극의 경우, 기준 평가치보다 더 비유창하게 지각된다면 100점보다 큰 점수를, 덜 비유창하게 지각된다면 100점보다 작은 점수를 매기도록 하였다. 기본적으로 배점은 음수나 0점은 허용하지 않으며 1점 이상 무한대로만 하도록 하였다. DME 평가에서는 평가하는 연장음 자극을 제시하기 전에 하나의 동일한 평가 기준치(1,005 ms)를 모두 들려주었다. 즉, 평가 기준치와 이후 평가해야 하는 음성 샘플을 1:1의 비율로 제시하였다(McHenry, 1999). 이를 통해 평가 기준치 대비 연장의 길이에 따른 비유창성 정도 평가에 있어 평가자의 ‘내적 기준’(internal standard)의 변동을 최소화하고자 했다(Eadie & Doyle, 2005b). EAI와 DME 평가에서 실험자극의 제시는 이월효과(carryover effect)를 고려해 개별 실험대상자에 따라 다른 순서로 제시하였다. 본 실험 전 실험대상자에게 각각 EAI와 DME 평가 방식에 대한 충분한 설명을 제공하였으며 각각의 본 실험 평가에서의 수행의 용이성을 위해 각각 5개의 연습 문항을 제공하였다. 그림 2와 그림 3은 본 연구에서 사용된 EAI와 DME의 평가 예시이다.

음성을 듣고 답해 주세요. [1~25번]
모두 정답이 없는 문항입니다.

1. 화자의 음성을 듣고 어떻게 느끼셨습니까? *

- ① 정상
- ② 미비(minimal)
- ③ 경도(mild)
- ④ 경중도(mild to moderate)
- ⑤ 중도(moderate)
- ⑥ 중심도(moderate to severe)
- ⑦ 심도(severe)

2. 화자의 음성을 듣고 어떻게 느끼셨습니까? *

- ① 정상
- ② 미비(minimal)
- ③ 경도(mild)
- ④ 경중도(mild to moderate)
- ⑤ 중도(moderate)
- ⑥ 중심도(moderate to severe)
- ⑦ 심도(severe)

EAI, equal-appearing interval scale.

그림 2. EAI(1-7점) 청지각적 평가 예시
Figure 2. A sample of EAI scale used for the auditory perceptual experiment

먼저 청취하신 기준 음성이 100점이라면, 1번 음성에 몇 점을 주시겠습니까? *

평가 기준 : 기준 음성보다 길면 그 정도에 따라 100점을 초과하는 점수를 기입하시고, 기준 음성보다 짧다면 그 정도에 따라 100점 미만의 점수를 기입하시면 됩니다. 점수 범위는 1점 이상 ~ 무한대 (단, 0점은 불가)

먼저 청취하신 기준 음성이 100점이라면, 2번 음성에 몇 점을 주시겠습니까? *

평가 기준 : 기준 음성보다 길면 그 정도에 따라 100점을 초과하는 점수를 기입하시고, 기준 음성보다 짧다면 그 정도에 따라 100점 미만의 점수를 기입하시면 됩니다. 점수 범위는 1점 이상 ~ 무한대 (단, 0점은 불가)

먼저 청취하신 기준 음성이 100점이라면, 3번 음성에 몇 점을 주시겠습니까? *

평가 기준 : 기준 음성보다 길면 그 정도에 따라 100점을 초과하는 점수를 기입하시고, 기준 음성보다 짧다면 그 정도에 따라 100점 미만의 점수를 기입하시면 됩니다. 점수 범위는 1점 이상 ~ 무한대 (단, 0점은 불가)

먼저 청취하신 기준 음성이 100점이라면, 4번 음성에 몇 점을 주시겠습니까? *

평가 기준 : 기준 음성보다 길면 그 정도에 따라 100점을 초과하는 점수를 기입하시고, 기준 음성보다 짧다면 그 정도에 따라 100점 미만의 점수를 기입하시면 됩니다. 점수 범위는 1점 이상 ~ 무한대 (단, 0점은 불가)

DME, direct magnitude estimation.

그림 3. DME 청지각적 평가 예시
Figure 3. A sample of DME scale used for the auditory perceptual experiment

2.4. 자료분석

음성 샘플에 대한 평가자의 반응 평균값을 구하였다. 즉, EAI의 경우에는 산술평균(arithmetic mean)을, DME는 기하평균(geometric mean)을 구하였다. DME 결과는 평가자마다 평가 기준치 대비 평가하는 연장음 음성 샘플의 비유창성 정도를 나타내는 숫자에 대한 개념이 서로 다르고 일부 평가자의 경우에는 다른 평가자에 비해 상대적으로 큰 숫자를 선택할 가능성이 있기에 평가 결과의 중앙치를 대표하는 값으로 기하평균을 구하였다(Engen, 1971; Ha, 2009; Stevens, 1974). 이후 EAI의 산술평균과 DME의 기하평균의 산포도를 작성한 후, 모형 분석을 통해 두 측정치 간의 관계가 선형적(linear)인지 아니면 비선형적(curvilinear)인지 R^2 값을 통해 조사하였다. 통계분석은 IBM SPSS Statistics for Windows(Version 27; IBM, Armonk, NY, USA)를 이용하였다.

2.5. 신뢰도 분석

EAI와 DME에서 실험대상자의 전체 평가 결과치에 대한 평가자 내(intrarater)와 평가자 간(interrater) 신뢰도를 절대합치도(absolute agreement)로 계산하였다. 결과적으로 평가자 내 신뢰도는 EAI가 DME보다 높게 나타났으나 평가자 간 신뢰도는 EAI가 DME보다 낮게 나타났다. 표 1에서 각각의 신뢰도 분석 결과를 제시하고 있다.

표 1. 신뢰도 분석 결과
Table 1. Results of reliability test

Intrarater	ICC	95% CI	p-value
EAI	0.947	(0.915, 0.970)	<0.001
DME	0.801	(0.668, 0.899)	<0.001
Interrater	ICC	95% CI	p-value
EAI	0.891	(0.817, 0.944)	<0.001
DME	0.916	(0.859, 0.958)	<0.001

ICC, intra-class correlation coefficient; CI, confidence interval; EAI, equal-appearing interval scale; DME, direct magnitude estimation.

3. 연구결과

3.1. 등간척도법(Equal-Appearing Interval Scale, EAI)과 직접크기평정법(Direct Magnitude Estimation, DME)에 의한 평균 결과

연장음의 길이에 따른 비유창성 지각 정도에 대한 청지각적 평가에서 EAI와 DME를 통한 평가 결과를 분석하였으며, 각각 EAI의 평가 결과는 산술평균으로 DME의 평가 결과는 기하평균으로 산출하였다. 연장음 길이에 따른 청지각적 평가 결과는 표 2에 제시하였다.

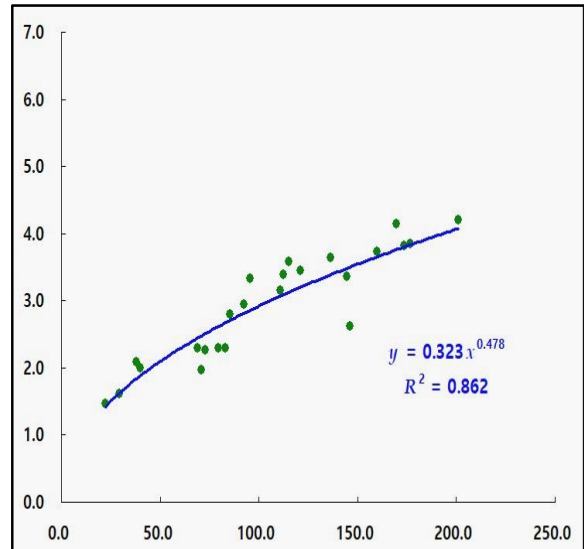
표 2. 연장 길이에 따른 청지각적 평가 결과
Table 2. Results of the auditory perceptual experiment

Stimulus (prolonged length in a unit of ms)	Arithmetic mean for EAI	Geometric mean for DME
285 (80)	1.47	22.25
365 (160)	1.62	29.51
445 (240)	2.00	40.12
525 (320)	2.09	38.19
605 (400)	1.97	71.03
685 (480)	2.29	69.10
765 (560)	2.26	72.75
845 (640)	2.29	82.92
925 (720)	2.29	79.41
1,085 (800)	2.79	85.48
1,165 (960)	2.94	92.59
1,245 (1,040)	3.32	95.60
1,325 (1,120)	3.38	112.52
1,405 (1,200)	3.15	110.76
1,485 (1,280)	3.59	115.17
1,565 (1,360)	3.44	121.17
1,645 (1,440)	3.65	136.5
1,725 (1,520)	2.62	146.06
1,805 (1,600)	3.74	159.72
1,885 (1,680)	3.35	144.65
1,965 (1,760)	3.85	176.74
2,045 (1,840)	3.82	173.42
2,125 (1,920)	4.15	169.79
2,205 (2,000)	4.21	201.05

1,005 ms의 경우 평가 기준치(modulus)로 사용되었기에 제외함.
EAI, equal-appearing interval scale; DME, direct magnitude estimation.

3.2. 등간척도법(Equal-Appearing Interval Scale, EAI)과 직접크기평정법(Direct Magnitude Estimation, DME)의 평가 결과치의 관계 분석 결과

두 평가 결과치의 산포도에 대한 모형 분석 결과, 선형 모형(linear model)에 의한 적합도는 $R^2=0.844(p\text{-value}<0.001)$, 이차 모형(quadratic model)에 의한 적합도는 $R^2=0.855(p\text{-value}<0.001)$, 거듭제곱 모형(power model)에 의한 적합도는 $R^2=0.862(p\text{-value}<0.001)$ 로 가장 높게 나타났다. 결국 EAI와 DME의 두 평가 측정치 간의 관계가 비선형적으로 나타났다. 연장음 길이에 따른 비유창성 지각 정도에 대한 청지각적 평가의 EAI와 DME 결과치에 대한 산포도는 그림 4에 제시하였다.



EAI, equal-appearing interval scale; DME, direct magnitude estimation.

그림 4. EAI와 DME 평가 결과치에 대한 산포도(X축: DME의 기하평균, Y축: EAI의 산술평균)

Figure 4. Scatter-plot of EAI and DME ratings for a perceived degree of disfluency based on prolonged length (X-axis: DME geometric means, Y-axis: EAI arithmetic means)

4. 논의와 결론

본 연구는 연장음의 길이에 따른 비유창성 지각 정도에 대해 각각 등간척도법(EAI)과 직접크기평정법(DME)을 통해 청지각적 평가를 실시한 후, 두 평가의 결과치의 관계가 선형적인(linear) 또는 비선형적인(curvilinear) 관계를 보이는지를 알아보고자 진행되었다. 이를 통해 결국 연장음의 길이에 따른 비유창성 지각 정도에 대한 적절한 평가 방법을 제안하고자 하였다. 산포도의 모형 분석 결과, 두 평가 결과치의 관계가 비선형적인 것으로 나타나 EAI보다는 DME가 적절한 평가 방법인 것으로 나타났다. 이는 연장음 길이에 따른 비유창성 정도에 대한 지각 속성이 양적 연속성의 특성임을 보여주는 결과라 할 수 있다. 이러한 심리물리적인 속성으로 인해 평가자가 연장음의 길이에 따른 비유창성 지각 정도를 EAI에서처럼 동일한 간격으로 분할하는 데 어려움을 보일 수 있다. 결론적으로 본 연구의 결과를 통해 연장음의 길이에 따른 비유창성 지각 정도를 청지각적으로 평가함에 있어 EAI보다는 DME가 타당한 평가방법으로 제안하고자 한다.

본 연구결과를 바탕으로 몇 가지 논의를 하자면, 첫째, 본 연구의 결과를 통해 연장음의 길이에 따른 비유창성 지각 정도에 있어 EAI보다는 DME가 적절한 평가 방법임을 알 수 있었다. 특히, 본 연구 결과는 말더듬의 심각 정도에 대한 청지각적 평가를 수행한 Schiavetti et al.(1983)의 결과와 일치한다. 서론에서 기술한 것처럼 Schiavetti et al.(1983)의 연구에서는 주로 막힘과 반복이 관찰되는 말더듬 성인을 대상으로 각각 EAI와 DME 평가를 수행하였으며, 두 평가 결과

치가 비선형적인 관계로 나타나면서 DME가 적절한 평가법이라고 제안하였다. 연장은 ‘비운율적 발성’(disrhythmic phonation)의 한 형태로 반복과는 질적으로 구별되는 말더듬 핵심행동이다(Yairi, 1990). 또한 성인 말더듬의 경우 연장은 상대적으로 적은 빈도수로 관찰된다(Jeon & Jeon, 2015). 본 연구는 기본적으로 연장의 경우, 주로 막힘과 반복이 관찰되는 말더듬 성인의 심각 정도를 평가한 Schiavetti et al.(1983)과는 다른 결과가 나타나지 않을까 하는 가능성에 대한 질문에서 진행되었다. 하지만 본 연구 결과를 통해 연장음의 길이에 따른 비유창성 지각 정도에 대한 평가도 전반적인 말더듬의 심각 정도와 같이 DME가 적절한 평가 방법이라는 답변을 얻을 수 있었다.

둘째, 기관식도 발성자의 음성의 용인도(acceptability)를 연구한 Eadie & Doyle(2005a)에서는 EAI와 DME의 두 평가 결과치의 관계가 선형적으로 결국 두 평가 모두 타당한 평가법이라고 제안하였다. 더불어 이들은 기관식도 발성자의 음성의 듣기 좋은 정도(pleasantness)에 대해서도 연구를 진행했는데 음성의 용인도와는 다른 결과가 나타났다. 즉, 음성의 듣기 좋은 정도는 두 평가 결과치의 관계가 비선형적으로 나타나면서 EAI보다는 DME가 적절한 평가법이라는 결론에 이른 것이다. 하지만 이러한 결과에 대해 연구자들은 음성의 듣기 좋은 정도에 대해 비선형 모델로 설명되는 변량과 선형 모델로 설명되는 변량과의 차이가 크지 않고, 음성의 용인도와 음성의 듣기 좋은 정도가 유의미한 상관관계를 보이며, DME보다는 EAI가 사용에 편리하기에 두 가지 음성평가에서는 EAI를 사용해도 무방하다고 제안하였다. 본 연구 결과에서 보는 것처럼(연구결과 3.2. 참조) 연장음의 길이에 따른 비유창성 지각 정도에 대해 비선형 모델로 설명되는 변량($R^2=0.862$)과 선형 모델로 설명되는 변량($R^2=0.844$)과의 차이가 크지 않다. 또한 비선형모형과 더불어 선형모형도 통계적으로 유의한 결과를 보이고 있다(p -value<0.001). 실제 임상에서의 수행의 편리성까지 고려할 때 EAI를 통한 평가도 가능하지 않을까 조심스럽게 제안해본다. 향후 통계적으로 좀 더 신뢰로운 결과를 도출하기 위해서 보다 많은 수의 평가자를 대상으로 실험을 진행할 필요가 있다고 하겠다.

셋째, 본 연구는 한국어를 모국어로 하는 화자를 대상으로 평마찰음 /s/의 연장음의 길이에 따른 비유창성 지각 정도에 대한 적절한 평가 방법에 대한 주제로 진행하였다. 하지만 연장은 단순히 정상 화자의 경우(Gósy & Eklund, 2018) 뿐 아니라 병리적인 유창성 문제를 보이는 말더듬 화자의 경우에도 나타난다. 말을 더듬는 사람들에게 나타나는 연장은 (정상 화자보다는) 더 길게 또는 시각적 긴장(tension)을 동반하면서 나타난다(Gregory, 2003). 이런 맥락에서 실제 말더듬 화자에게 나타나는 연장에 대한 청지각적 평가 연구도 진행해 볼 필요가 있겠다. 이를 위해 말더듬 진단 및 중재 전문가인 언어재활사를 대상으로 단순히 음성샘플을 통한 평가를 넘어 시청각적인(audiovisual) 샘플을 통한 연구를 진

행해 볼 것을 제안한다.

본 연구의 제한점과 관련 향후 연구에 대해 기술하자면, 먼저 비교적 적은 수의 평가자가 참여하였다는 것이다. 향후 통계적으로 좀 더 신뢰로운 결과를 도출하기 위해 보다 많은 수의 평가자를 대상으로 실험을 진행할 필요가 있겠다. 또한 본 연구에서는 한국어를 모국어로 사용하는 일반 화자를 대상으로 하여 연장음의 길이에 따른 비유창성 지각 정도에 적절한 평가 방법을 제안하였다. 하지만 유창성장애의 진단 및 중재 전문가인 언어재활사를 대상으로 한 비교 연구도 진행해 볼 필요가 있다. 연장음의 길이에 따른 비유창성 지각 정도에 있어 비전문가와 전문가의 청지각적 차이가 존재할 가능성이 있기 때문이다(Park et al., 2018). 마지막으로 ‘들을 수 없는 연장’(inaudible prolongation)(Van Riper, 1982)인 막힘의 길이에 따른 비유창성 지각 정도에 대해서도 EAI와 DME의 비교 연구를 통한 적절한 평가 방법을 규명하는 연구도 제안해 본다. 말더듬 비유창성의 가장 진전된 형태로서 막힘에 대한 적절한 평가 방법에 대한 규명도 분명 흥미로운 연구 주제가 될 수 있기 때문이다.

감사의 글

본 연구에 평가자로 참여해주신 모든 대상자에게 감사드립니다.

References

- Darley, F. L., Aronson, A. E., & Brown, J. R. (1969). Differential diagnostic patterns of dysarthria. *Journal of Speech and Hearing Research, 12*(2), 246-269.
- Duffy, J. R. (1995). *Motor speech disorders: Substrates, differential diagnosis, and management*. St. Louis, MO: Elsevier Mosby.
- Eadie, T. L., & Doyle, P. C. (2002). Direct magnitude estimation and interval scaling of pleasantness and severity in dysphonic and normal speakers. *Journal of the Acoustical Society of America, 112*(6), 3014-3021.
- Eadie, T. L., & Doyle, P. C. (2005a). Scaling of voice pleasantness and acceptability in tracheoesophageal speakers. *Journal of Voice, 19*(3), 373-383.
- Eadie, T. L., & Doyle, P. C. (2005b). Classification of dysphonic voice: Acoustic and auditory-perceptual measures. *Journal of Voice, 19*(1), 1-14.
- Engen, T. (1971). Psychophysics: II. Scaling methods. In R. S. Woodworth, J. W. Kling, H. Schlosberg, & L. Riggs (Eds.), *Woodworth & Schlosberg's experimental psychology* (pp. 47-86). New York, NY: Holt, Rinehart, and Winston.
- Gósy, M., & Eklund, R. (2018). Language-specific patterns of segment prolongation in Hungarian. *The Phonetician, 115*,

- 36-52.
- Gregory, H. H. (2003). *Stuttering therapy: Rationale and procedures*. Boston, MA: Allyn & Bacon.
- Ha, S. (2009). A comparison of equal-appearing interval scaling and direct magnitude estimation in the perceptual judgment of hypernasality. *Communication Sciences & Disorders, 14*(4), 563-573.
- Hoodin, R. B., & Gilbert, H. R. (1989). Nasal airflows in Parkinsonian speakers. *Journal of Communication Disorders, 22*(3), 169-180.
- Jeon, H. S., & Jeon, H. E. (2015). Characteristics of disfluency clusters in adults who stutter. *Journal of Speech-Language & Hearing Disorders, 24*(1), 135-144.
- Jones, K., Logan, K. J., & Shrivastav, R. (2005, November). Duration, rate, and phoneme-type effects on listeners' judgments of prolongations. *Proceedings of the Annual Meeting of the American Speech-Language-Hearing Association*. San Diego, CA.
- Karnell, M. P., Folkins, J. W., & Morris, H. L. (1985). Relationships between the perception of nasalization and speech movements in speakers with cleft palate. *Journal of Speech, Language, and Hearing Research, 28*(1), 63-72.
- Kawai, N., Healey, E. C., & Carrell, T. D. (2007). Listeners' identification and discrimination of digitally manipulated sounds as prolongations. *The Journal of the Acoustical Society of America, 122*(2), 1102-1110.
- Kim, H. H. (1996, February). Perceptual, acoustical, and physiological tools in ataxic dysarthria management: A case report. *Proceedings of the Korean Society of Phonetic Sciences and Speech Technology Conference* (pp. 9-22). Seoul, Korea.
- Kreiman, J., & Gerratt, B. R. (1998). Validity of rating scale measures of voice quality. *Journal of the Acoustical Society of America, 104*(3), 1598-1608.
- Lee, K. H. (2001). *A study of Korean lenis fricatives* (Doctoral dissertation). Korea University, Seoul, Korea.
- Lee, Y., Park, H., Lim, D., & Kim, G. (2022). Usefulness of direct magnitude estimation (DME) in auditory perceptual assessments measuring dysphonia severity. *Journal of Voice*. <https://doi.org/10.1016/j.jvoice.2022.09.001>
- Ludlow, C. L., & Bassich, C. J. (1984). Relationship between perceptual ratings and acoustic measures of hypokinetic speech. In: M. McNeil, J. C. Rosenbek, & A. E. Aronson (Eds.), *The dysarthrias: Physiology, acoustic, perception, management* (pp. 163-196). San Diego, CA: College Hill Press.
- Martin, R. (1965). Direct magnitude estimation judgments of stuttering severity using audible and audible-visible speech samples. *Speech Monographs, 32*(2), 169-177.
- Martin, R. R., Haroldson, S. K., & Triden, K. A. (1984). Stuttering and speech naturalness. *Journal of Speech and Hearing Disorders, 49*(1), 53-58.
- McHenry, M. A. (1999). Aerodynamic, acoustic, and perceptual measures of nasality following traumatic brain injury. *Brain Injury, 13*(4), 281-290.
- Metz, D. E., Schiavetti, N., & Sacco, P. R. (1990). Acoustic and psychophysical dimensions of the perceived speech naturalness of nonstutterers and posttreatment stutterers. *Journal of Speech and Hearing Disorders, 55*(3), 516-525.
- Park, J., & Chung, I. (2023). Korean listeners' identification and discrimination of lengthened /s/ as prolongations. *Clinical Linguistics & Phonetics, 13*, 1-17.
- Park, J., Jun, J. P., & Chung, I. (2018). Comparison of perception of the prolonged /s/ in Korean by average adult listeners and speech-language pathologists. *Audiology and Speech Research, 14*(3), 184-193.
- Prather, E. M. (1960). Scaling defectiveness of articulation by direct magnitude-estimation. *Journal of Speech and Hearing Research, 3*(4), 380-392.
- Price, D. D., Staud, R., & Robinson, M. E. (2012). How should we use the visual analogue scale (VAS) in rehabilitation outcomes? II: Visual analogue scales as ratio scales: An alternative to the view of Kersten et al. *Journal of Rehabilitation Medicine, 44*(9), 800-801.
- Schiavetti, N. (1992). Scaling procedures for the measurement of speech intelligibility. In R. D. Kent (Ed.), *Intelligibility in speech disorders: Theory, measurement and management* (pp. 11-34). Amsterdam, The Netherlands: John Benjamins.
- Schiavetti, N., Martin, R. R., Haroldson, S. K., & Metz, D. E. (1994). Psychophysical analysis of audiovisual judgments of speech naturalness of nonstutterers and stutterers. *Journal of Speech, Hearing Research, 37*(1), 46-52.
- Schiavetti, N., Metz, D. E., & Sitler, R. W. (1981). Construct validity of direct magnitude estimation and interval scaling of speech intelligibility: Evidence from a study of the hearing impaired. *Journal of Speech, Language, and Hearing Research, 24*(3), 441-445.
- Schiavetti, N., Sacco, P. R., Metz, D. E., & Sitler, R. W. (1983). Direct magnitude estimation and interval scaling of stuttering severity. *Journal of Speech, Language, and Hearing Research, 26*(4), 568-573.
- Sewall, A., Weglarski, A., Metz, D. E., Schiavetti, N., & Whitehead, R. (1999). A methodological control study of scaled vocal breathiness measurements. *Contemporary Issues in Communication Science and Disorders, 26*, 168-172.
- Southwood, M. H. (1996). Direct magnitude estimation and interval scaling of naturalness and bizarreness of the dysarthria

- associated with amyotrophic lateral sclerosis. *Journal of Medical Speech-Language Pathology*, 4(1), 13-25.
- Stevens, S. S. (1974). *Psychophysics: Introduction to its perceptual, neural, and social prospects*. New York, NY: Wiley.
- Toner, M. A., & Emanuel, F. W. (1989). Direct magnitude estimation and equal appearing interval scaling of vowel roughness. *Journal of Speech, Language, and Hearing Research*, 32(1), 78-82.
- Van Riper, C. (1982). *The nature of stuttering*. Englewood Cliffs, NJ: Prentice-Hall.
- Weismer, G., & Laures, J. S. (2002). Direct magnitude estimates of speech intelligibility in dysarthria: Effects of a chosen standard. *Journal of Speech, Language, and Hearing Research*, 45(3), 421-433.
- Whitehill, T. L., Lee, A. S. Y., & Chun, J. C. (2002). Direct magnitude estimation and interval scaling of hypernasality. *Journal of Speech, Language, and Hearing Research*, 45(1), 80-88.
- Yairi, E. (1990). Subtyping child stutterers for research purpose. In J. A. Cooper (Ed.), *Research needs in stuttering: Roadblocks and future directions* (pp. 50-57). Rockville, MD: American Speech-Language-Hearing Association.
- Zraick, R. I., & Liss, J. M. (2000). A comparison of equal-appearing interval scaling and direct magnitude estimation of nasal voice quality. *Journal of Speech, Language, and Hearing Research*, 43(4), 979-988.

• **박진 (Jin Park)** 교신저자

가톨릭관동대학교 언어재활학과 조교수
강원특별자치도 강릉시 범일로 579번길 24
Tel: 033-649-7737
Email: gatorade70@cku.ac.kr
관심분야: 유창성장애, 음성장애

• **차화정 (Hwajung Cha)**

가톨릭관동대학교 언어재활학과 석사과정
강원특별자치도 강릉시 범일로 579번길 24
Tel: 033-649-7670
Email: slpcha@kakao.com
관심분야: 유창성장애, 언어발달장애

• **배세진 (Sejin Bae)**

가톨릭관동대학교 언어재활학과 석사과정
강원특별자치도 강릉시 범일로 579번길 24
Tel: 033-649-7670
Email: baesejin@cku.ac.kr
관심분야: 신경언어장애, 보완대체의사소통

연장음 길이에 따른 비유창성 정도 평가: 등간척도와 직접크기평정 비교 연구*

박진·차화정·배세진
가톨릭관동대학교 언어재활학과

국문초록

본 연구는 연장음의 길이에 따른 비유창성 지각 정도에 대해 각각 등간척도와 직접크기평정을 통한 청지각적 평가를 실시한 후, 두 평가의 결과치가 선형적인 또는 비선형적인 관계를 보이는지를 알아보고자 진행되었다. 이를 통해 연장음의 길이에 따른 비유창성 지각 정도에 대한 적절한 평가 방법을 제안하고자 하였다. 이를 위해 한국어를 모국어로 하는 만 19세 이상 성인 남녀 34명(남: 9명, 여: 25명, 평균연령: 32.9세)이 평가자로 참여하였다. 실험참여자 는 먼저 한국어 평마찰음 /s/를 원래 길이에서 80 ms씩 연장하여 2,000 ms(i.e., 285 ms, 365 ms., ..., 2,125 ms, 2,205 ms) 까지 연장 변조한 총 25개의 자극이 들어 있는 문장을 듣고, 등간척도(1-7점, 1은 '정상', 7은 '심도')로 평가하였다. 이후에 등간척도 평가 결과, '경중도'(4점)에 해당하는 음성샘플을 선정해 이를 기준 평가치(modulus)로 하여 직접 크기평정을 실시하였다. 두 평가 결과치에 대한 산포도를 작성한 후, 모형 분석을 통해 두 측정치 간의 관계가 선형적(linear)인지 비선형적(curvilinear)인지 R^2 값을 통해 조사하였다. 연구 결과, 두 평가 결과치의 관계가 비선형적인 양상을 보이는 것으로 나타났으며 이는 연장음의 길이에 따른 비유창성 정도 평가에 있어 등간척도보다는 직접크기평정이 적절한 평가 방법임을 보여주는 결과이다.

핵심어: 연장음, 등간척도, 직접크기평정, 청지각적 평가, 말더듬(비유창성)

참고문헌

- 김향희(1996). 운동실조형 마비성 구음장애에 적용되는 지각적, 음향학적, 생리학적 도구에 관하여: 환자사례를 중심으로. *제 2 회 음성학 학술대회자료집* (pp. 68-71).
- 이경희(2001). *국어 마찰음 연구*. 고려대학교 박사학위논문.
- 전희숙, 전효은(2015). 말더듬 성인의 비유창성군 특성. *언어치료 연구*, 24(1), 135-144.
- 하승희(2009). 과대비성의 청지각적 평가의 등간 척도와 직접 크 기 측정법 비교 연구. *언어청각장애연구*, 14(4), 563-573.

* 본 과제(결과물)는 2023년도 교육부의 재원으로 한국연구재단의 지원을 받아 수행된 지자체-대학 협력기반 지역혁신 사업의 결과입니다(2022 RIS-005).