Speech-Perceptual Map Using Confusing Matrix and Application of Shepard's Law

Chanbeom Kwak^{1,2}, Sunghwa You^{1,2}, Woojae Han^{1,2,3}, Donguk Lee^{1,2,4},
Sunmi Ma^{1,2}, Sihoon Park^{1,2}, Yeju Kim^{1,2}

¹Laboratory of Hearing and Technology, ²Division of Speech Pathology and Audiology, ³Research Institute of Audiology and Speech Pathology, College of Natural Sciences, Hallym University, Chuncheon, Korea

⁴Department of Audiology and Speech Pathology, The University of Tennessee Health Science Center, Knoxville, TN, USA

오차행렬과 Shepard's 법칙을 적용한 말지각 지도

곽찬범^{1,2} · 유성화^{1,2} · 한우재^{1,2,3} · 이동욱^{1,2,4} · 마선미^{1,2} · 박시훈^{1,2} · 김여주^{1,2}

한림대학교 자연과학대학 HearT 실험실¹, 언어청각학부², 청각언어연구소³, 테네시대학교 청각언어병리학과 건강과학센터⁴

Purpose: To discriminate and/or perceive phonemes is one of the most important abilities in human communication. The purpose of current study was to measure the confusing matrix, similarity, and distance of Korean phonemes based on Shepard's law and to provide Korean speech-perceptual map. **Methods:** To measure the similarity and distance, twenty young listeners with normal hearing conducted the speech discrimination task, which randomly assigned consonant-vowel (CV) combinations of 18 Korean consonants with /a/ vowel under quiet environment and two signal-to-noise ratio (SNR) noisy conditions (i.e., +3 and 0 dB SNR). The data of each participant was analyzed by classic confusion matrix. Based on the data of confusion matrix, the similarity and distance were calculated using the Shepard's law. After calculating the distance, the speech-perceptual map was designated. **Results:** In general, the CV similarity was increased as a level of the background noise increased. The CV distance was also decreased as the noise levels increased. In quiet condition, only 18 pairs were showed confusing to discriminate. For the most confusing CV pair was /p/-/t/ which showed similarity of 0.14. With background noise, however, the +3 and 0 dB SNR conditions showed total 61 pairs and 64 pairs, respectively. The /p^h/-/t^h/ was most confusing pair in both +3 (1.00 for similarity) and 0 dB SNR (0.98 for similarity). **Conclusion:** The results of current study suggest that the development of speech-perceptual map should be considered in the clinical settings in terms of rehabilitation for the hearing impaired listeners.

Key Words: Phonemic perceptual similarity, Phonemic perceptual distance, Confusing phonemes, Speech perception in noise.

Received: March 25, 2019 / **Revised:** April 2, 2019 / **Accepted:** April 2, 2019

Correspondence: Woojae Han, Laboratory of Hearing and Technology, Division of Speech Pathology and Audiology, Research Institute of Audiology and Speech Pathology, # 8603 Natural Science Building, Hallym University, 1 Hallymdaehak-gil, Chuncheon 24252, Korea

Tel: +82-33-248-2216 / **Fax:** +82-33-256-3420 / **E-mail:** woojaehan@hallym.ac.kr

INTRODUCTION

음성언어를 통한 의사소통 능력은 교육과 재활이라는 측면에서 그 중요성이 강조되며(Tye-Murray, 2014) 이를 확인하고 평가하기 위한 방법으로 임상에서는 순음과 어음을 활용한 검사도구가 사용된다. 그러나 이러한 검사들은 조용한 환경에서의 가청력(audibility)을 평가하는 데 목적을 두고 있으며, 이는

일상생활에서 필연적으로 발생하는 배경소음 환경에서의 의사소통 능력을 예측하기 어렵다는 한계점을 지닌다(An et al., 2002). 또한 난청인의 경우에는 보청기, 인공와우 등 다양한 보장구의 착용을 통해 향상된 의사소통 능력을 기대할 수 있으나(Lee, 2016), 이는 음소(phoneme)를 명확하게 분별하거나 인지로 인한 결과가 아닌 가청력(audibility)의 향상으로 인한 결과로 분석된다(Ching et al., 2001; Firszt et al., 2004). 일반적으로 난청인은 건청인에 비해 상대적으로 저조한 음소 분별 및 인지 능력을 보이며 보청기, 인공와우 등의 보장구를 착용하더라도 배경소음 환경에서는 여전히 어려움을 지닌다(Killion, 1997;

© This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<https://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.

Lee & Han, 2019). 또한 난청인은 건청인에 비해 배경소음 환경에서의 어음 인지 및 분별 능력은 더욱 저조하며(Wilson & Strouse, 2002) 배경소음의 강도에 따라 달라진다(Plomp, 1986). 따라서 배경소음 유무에 따른 난청인의 보다 세밀하고 객관화된 어음 인지 및 분별 능력의 평가는 일상생활에서 이들이 느끼는 어려움을 명확하게 확인하고 착용한 보청구에 대한 실질적인 이득(benefit)을 제공하고, 더 나아가서는 청능재활 시의 평가와 재활에 대한 중요한 정보를 제공하게 된다(Beattie, 1989).

유사한 음소의 분별 및 인지는 타인과의 의사소통이라는 관점에서 중요한 요인이며 Miller & Nicely(1995)는 이를 오차행렬(confusing matrix)을 통해 확인하였다. 오차행렬은 각 행과 열에 자극음인 음소를 순차적으로 배치한 뒤 대상자의 반응에 따른 정/오반응을 기입하는 방식으로써, 대상자의 정반응은 대각선으로 나타나며 오반응은 대각선의 반응을 제외한 모든 반응을 의미한다. 이를 바탕으로 Shepard(1972)는 음소 자극에 대한 오차행렬에 유사성(similarity)과 거리(distance)의 개념을 적용한 Shepard's law를 통해 음소 간 분별이나 인지 시 가장 큰 오차를 보이는 음소를 확인하고자 하였다. 즉, 유사성(similarity)이란 오차행렬에서 나타난 오반응의 총합을 정반응의 총합으로 나눈 값을 의미하며, 거리(distance)는 오반응에 대한 개념을 다차원적인 척도로 표현하는 것을 의미한다(Shepard, 1972). 예를 들어 Miller & Nicely(1955)의 연구에서는 발성 위치에 따른 자음 분별 시 시각적 인지와 청각적 인지는 다른 양상을 보이게 되며 영어에서는 /t/, /s/, /d/, /z/, /n/의 5개 자음을 청각적으로 분별하는 데 있어 가장 많은 혼동을 보였다. Shepard(1972)의 연구에서는 기존 선행연구의 결과를 Shepard's law를 적용하여 모델적합도(goodness of fit)를 확인하였고 99.4%의 높은 적합도를 보고하였다. 이는 오차행렬에 기반한 유사성과 거리 분석이 대상자의 반응을 명확하게 표현할 수 있는 방법임을 지지하였다. Johnson(2003)의 저서에서는 Shepard's law를 적용하여 혼동되는 음소에 대한 말지각 지도(speech-perceptual map)에 대한 접근법을 소개한 바 있다. 즉 위치를 나타내는 지도의 개념을 토대로 말지각적인 측면에서도 각 음소에 대한 거리를 계산하여 음소별 혼동되는 정도를 직관적으로 확인할 수 있는 방법이다. 최근까지도 음소 간 유사성과 거리에 대한 음향학적 모델 혹은 말지각 지도를 제작하려는 다양한 연구들이 진행되고 있으나 아직까지 한국어 기반의 음소 간 유사성과 거리에 대한 연구는 부재한 실정이다. 이에 본 연구에서는 Shepard's law를 기반으로 한국어 자모음의 오차행렬과 유사성, 거리를 측정하고 이를 활용하여 한국어의 말지각 지도를 제시하고자 한다.

MATERIALS AND METHODS

연구 대상

정상 청력을 가진 20대 성인 20명(남녀 각각 10명; 평균 연령: 22.05세, 표준 편차: ± 1.43)이 본 연구에 참여하였다. 연구 대상자는 과거 이과적 병력이 없고, 고막운동성 검사에서는 A 유형을 나타냈다. 또한 0.5, 1, 2, 4 kHz의 4분법 순음평균역치는 우측 1.94 dB HL, 좌측 2.56 dB HL이었으며, 기도-골도 차는 양측 귀 모두 2.13 dB HL로 모든 대상자들은 20 dB HL 이하의 정상 청력 기준을 충족하였다. 실험에 참여하기 전 대상자들은 연구의 목적과 방법에 대하여 자세한 설명을 듣고 이해한 후 참여 동의서에 서명을 하였다.

자극음 및 연구 절차

본 연구에서는 한국어의 18개 자음과 1개의 /a/ 모음으로 결합된 총 18개의 자모음을 사용하였다. 구체적으로 자음은 /p/, /p^{*}/, /p^h/, /t/, /t^{*}/, /t^h/, /k/, /k^{*}/, /k^h/, /ts/, /ts^{*}/, /ts^h/, /s/, /s^{*}/, /m/, /n/, /l/, /h/였다. 조음 위치에 따른 분류로 /p/, /p^{*}/, /p^h/, /m/는 두 입술을 사용하여 조음하는 양순음이었으며, /t/, /t^{*}/, /t^h/, /n/는 혀끝이나 혀날을 윗잇몸에 붙이거나 간격을 좁혀 조음하는 치조음이었다. /k/, /k^{*}/, /k^h/는 혀의 뒷부분과 연구개 사이에서 조음하는 연구개음이었으며, /ts/, /ts^{*}/, /ts^h/는 혀의 앞부분과 경구개 사이에서 조음하는 경구개음이었다. 또한, /s/, /s^{*}/, /l/와 /h/는 각각 치경음과 성문음이었다(Kim, 2002).

자극음은 총 4명의 화자(남녀 각각 2명)가 발화한 자모음을 음성분석 프로그램인 Computerized Speech Lab (CSL™; Kay-PENTAX, Montvale, NJ, USA)을 사용하여 마이크로 녹음하였고 Adobe Audition 프로그램(Ver. CC 2014.2; Adobe Systems Complex, San Jose, CA, USA)으로 자극음의 실효값(root-mean square)을 조절하였다. 이를 조용한 환경과 +3, 0 dB signal-to-noise ratio (SNR)의 2가지 백색잡음(white noise) 조건에 적용하여 검사를 실시하였다.

자극음 제시는 두 개의 스피커와 연결된 노트북(NT910S3G-K32B, Samsung, Seoul, Korea)과 Superlab 프로그램(version 5.0, Cedrus Corporation, San Pedro, CA, USA)을 통해 각 대상자가 원하는 크기의 최고 쾌적 수준(most comfortable level)에서 재생하였다. 두 개의 스피커를 대상자의 양 옆 각각 45° 방향에서 1 m 거리에 두었고 대상자는 의자에 앉아 제시되는 자극음을 듣고 들은 자극음과 가장 비슷한 자모음을 18개의 보기에서 고르도록 요청하였다. 실험 시작 전, 각 대상자에게는 들은 자극음에 대한 반응을 기입하는 오차행렬 기록지가 제공되었다. 실험은 배경소음 유무와 강도에 따라 쉬움(quiet), 보통(+3 dB SNR), 어려움(0 dB SNR)으로 구분되었으며 순차적으

로 진행되었다. 자극음은 각 난이도에서 1개의 자모음당 20번의 반복을 무작위(pseudo-randomized)로 제시하였으며 20명의 연구 대상자의 총 제시 횟수는 자모음 1개당 400회였다.

자료 분석

제시된 18개의 자모음에 대한 20명의 연구 대상자들의 응답은 18×18 의 오차행렬을 통해 자극음 제시 횟수와 대상자의 정/오반응 횟수에 따른 백분율로 계산되었다. 오차행렬에서 대각선상에 표기되는 반응은 정반응으로 간주하였으며, 대각선상에 표기되지 않은 모든 반응은 오반응으로 간주하였다. 정반응과 오반응에 대한 기준은 제시되는 자극음과 대상자가 인지한 자모음이 일치할 경우에는 정반응으로, 일치하지 않을 경우에는 오반응으로 처리하였다. 백분율로 계산한 오차행렬에서의 반응은 Shepard's law에 근거하여 말지각 유사성을 확인 및 분석하였다.

유사성(similarity, S_{ij})은 오반응을 보인 반응의 총합을 정반응을 보인 전체 반응의 총합으로 나누는 것으로 계산하였고[1], 유사성과 자극음의 혼동은 정비례 관계, 즉 유사성이 높을수록 자극음의 혼동 역시 높게 나타나며, 유사성이 낮을수록 자극음의 혼동은 낮게 나타난다.

$$S_{ij} = \frac{(P_{ij} + P_{ji})}{(P_{ii} + P_{jj})} \quad [1]$$

* S_{ij} 는 i 범주와 j 범주 간 유사성을 의미하고, P는 probability

를 뜻한다.

이후, 유사성 결과를 근거로 말지각 거리를 분석하였다. 거리(d_{ij})는 계산된 유사성에 음의 자연로그(-ln)를 취하였다[2]. 유사성과 자극음의 혼동 간의 관계와는 다르게, 거리는 자극음의 혼동과 반비례 관계, 즉 거리가 짧을수록 자극음의 혼동은 높게 나타나며, 거리가 멀수록 자극음의 혼동은 낮게 나타난다.

$$d_{ij} = -\ln(S_{ij}) \quad [2]$$

RESULTS

말지각 유사성

18개의 한국어 자모음에 따른 유사성 분석 결과, 소음이 증가함에 따라 음소 간 유사성의 빈도와 정도가 증가되는 것을 확인하였다. 조용한 환경에서는 /t/와 /p/ (0.14), /m/와 /p*/ (0.05), /p^h/와 /p/ (0.04) 등 총 17회(평균: 0.03, 표준편차: 0.03)의 유사성을 보였으나(Table 1), 배경소음이 제시된 +3 dB SNR 환경에서는 /t^h/와 /p^h/ (1.00), /ts/와 /t/ (0.70), /t/와 /p/ (0.62), /h/와 /p^h/ (0.55), /h/와 /t^h/ (0.50) 등 총 64회의 유사성(평균: 0.14, 표준편차: 0.20)을 보였으며 그 정도 역시 조용한 환경에 비해 증가되었다(Table 2). 0 dB SNR 환경에서는 /t^h/와 /p^h/ (0.98), /t/와 /p/ (0.88), /ts/와 /t/ (0.85), /h/와 /t^h/ (0.76), /s/와 /t^h/ (0.57), /h/와 /p^h/ (0.57) 등 총 61회의 유사성(평균: 0.18, 표준편차: 0.23)을 보였으며(Table 3), +3 dB SNR 환경과 비교

Table 1. Results of speech perceptual similarity from 20 normal hearing listeners at quiet condition

Stimuli	Responses																	
	p	p*	p ^h	t	t*	t ^h	k	k*	k ^h	ts	ts*	ts ^h	s	s*	m	n	l	h
p	1.00																	
p*	0.01	1.00																
p ^h	0.04		1.00															
t	0.14			1.00														
t*			0.01		1.00													
t ^h			0.03			1.00												
k	0.01			0.01			1.00											
k*								1.00										
k ^h							0.02		1.00									
ts										1.00								
ts*											1.00							
ts ^h												1.00						
s													0.02	1.00				
s*															1.00			
m		0.05														1.00		
n			0.00		0.05			0.00								0.00	1.00	
l																	1.00	
h	0.00		0.03															1.00

Table 2. Results of speech perceptual similarity from 20 normal hearing listeners at +3 dB SNR

Stimuli	Responses																	
	p	p*	p ^h	t	t*	t ^h	k	k*	k ^h	ts	ts*	ts ^h	s	s*	m	n	l	h
p	1.00																	
p*	0.01	1.00																
p ^h	0.03		1.00															
t	0.62		0.01	1.00														
t*		0.04		0.01	1.00													
t ^h			1.00	0.01	0.01	1.00												
k	0.26			0.26		0.01	1.00											
k*		0.08			0.10			1.00										
k ^h	0.00		0.12			0.20	0.01		1.00									
ts	0.23			0.70		0.01	0.11	0.03		1.00								
ts*					0.15			0.02			1.00							
ts ^h			0.17			0.31			0.07			1.00						
s	0.07		0.48	0.07		0.50			0.10	0.08		0.12	1.00					
s*		0.02			0.25			0.12			0.30			1.00				
m		0.07		0.00	0.01		0.00	0.01		0.01					1.00			
n				0.01	0.06			0.01		0.01	0.01			0.01	0.08	1.00		
l										0.01						0.03	1.00	
h	0.34	0.01	0.55	0.08		0.50	0.12		0.14	0.01		0.13	0.30					1.00

SNR: signal-to-noise ratio

Table 3. Results of speech perceptual similarity from 20 normal hearing listeners at 0 dB SNR

Stimuli	Responses																	
	p	p*	p ^h	t	t*	t ^h	k	k*	k ^h	ts	ts*	ts ^h	s	s*	m	n	l	h
p	1.00																	
p*	0.02	1.00																
p ^h	0.08		1.00															
t	0.88			1.00														
t*		0.08			1.00													
t ^h	0.03		0.98			1.00												
k	0.26		0.03	0.36		0.02	1.00											
k*		0.07			0.17			1.00										
k ^h			0.23			0.31	0.02		1.00									
ts	0.36		0.01	0.85		0.02	0.12			1.00								
ts*					0.15			0.03			1.00							
ts ^h			0.26		0.01	0.30			0.13			1.00						
s	0.13		0.43	0.16		0.57	0.01		0.19	0.09			1.00					
s*		0.02			0.22			0.17			0.37			1.00				
m	0.01	0.08	0.02					0.02							1.00			
n				0.02	0.05		0.01	0.02		0.01	0.00			0.00	0.11	1.00		
l								0.00		0.01						0.07	1.00	
h	0.27		0.57	0.06		0.76	0.16		0.29			0.12	0.30					1.00

SNR: signal-to-noise ratio

했을 때, 유사성의 발생 빈도는 미약한 감소를 보였으나, 평균적인 유사성의 정도는 증가함을 확인하였다.

말지각 지도

18개의 한국어 자모음에 따른 유사성 결과에 근거하여 말지

각 거리를 분석한 결과를 Figure 1로 제시하였다. 말지각 거리는 소음이 증가함에 따라 거리의 정도가 짧아지는 것을 확인하였으며 이는 유사성과 동일한 결과를 보였다. 조용한 환경에서는 /t/와 /p/ (1.96), /m/와 /p*/ (2.99) 등 평균 4.16 (표준편차: 1.12)의 거리를 나타냈다(Figure 1A). 배경소음이 제시된 +3

Figure 1. Speech perceptual-map for 18 Korean consonants as a function of SNR. A: Quiet. B: +3 dB SNR. C: 0 dB SNR. SNR = signal-to-noise ratio, d = distance.

dB SNR 환경에서는 /t^h/와 /p^h/ (-0.01), /ts/와 /t/ (0.36), /t/와 /p/ (0.48), /h/와 /p^h/ (0.59) 등 평균 2.97(표준편차: 1.59)의 거리를 보였다(Figure 1B). 0 dB SNR 환경에서는 /t^h/와 /p^h/ (0.02), /t/와 /p/ (0.13), /ts/와 /t/ (0.16), /h/와 /t^h/ (0.28), /h/와 /p^h/ (0.55) 등 평균 2.60(표준편차: 1.54)의 거리를 보였다(Figure 1C). 유사성 결과와 마찬가지로 거리 결과에서도 소음이 증가함에 따라 음소별 거리가 감소하는 것을 확인하였다.

DISCUSSIONS

본 연구는 20대의 정상 청력을 가진 성인 남녀를 대상으로 오차행렬과 Shepard's law에 근거하여 배경소음 정도에 따른 한국어 자모음의 유사성과 거리를 측정하고, 이를 기반으로 말 지각 지도를 제시하였다. 음소 간 유사성을 확인한 결과, 배경소음이 증가함에 따라 음소 간 유사성은 증가함을 확인하였다. 이는 배경소음이 증가함에 따라 정상 청력 성인의 어음인지 오류율이 증가함을 보고한 Chun et al.(2015)의 연구 결과가 뒷받침하고 있다. 구체적으로, /p/-/t/의 유사성은 조용한 환경, +3 dB SNR, 0 dB SNR에서 각각 0.14, 0.62, 0.88을 보였으며, 이는 배경소음이 증가함에 따라 동일한 음소의 분별 또는 인지에 대한 어려움이 증가함을 보인 것으로 확인하였다. Chun et al. (2015)의 연구에서는 폐쇄음과 마찰음이 배경소음에 관계없이 높은 오류율을 보인다고 보고한 바 있으며, 본 연구의 결과에서도 양순음에 해당하는 /p/와 치조음에 해당하는 /t/ 음소가 조음 위치에 따른 분류는 상이하나, 조음 방식에 따른 음소 분류에서 폐쇄음에 해당되기 때문이라고 생각한다. 또한, /t/-/ts/ 음소 간 유사성 결과에서도 +3과 0 dB SNR에서 각각 0.70과 0.85로 높은 유사성을 보인 것으로 확인하였으며 이는 치조음인 /t/ 음소가 폐쇄음에 해당되기 때문에 상대적으로 혼동되기 쉬운 음소이기 때문에 높은 유사성을 보인 것으로 분석하였다.

이는 배경소음이 제시됨에 따라 조음 방식뿐 아니라, 조음 위치에 따라라도 음소 간 혼동이 증가하는 것으로 보인다. 보다 자세히 두 가지 배경소음 환경에 대한 음소 간 유사성 결과에서는, /p^h/-/t^h/ (1.00, 0.98), /t/-/ts/ (0.70, 0.85), /p^h/-/h/ (0.55, 0.76), /t^h/-/s/ (0.50, 0.57), /t^h/-/h/ (0.50, 0.76), /p^h/-/s/ (0.48, 0.43)로 확인하였으며, 이는 높은 유사성을 보이는 음소 간 결과는 모두 폐쇄음 혹은 마찰음에 해당하는 음소가 포함되었기 때문으로 해석할 수 있다(Chun et al., 2015). 그러나 Han et al. (2017)의 연구에서는 폐쇄음에 비해 비음으로 인한 음소의 오류율이 더 높음을 보고한 바 있으나 이는 본 연구와의 자극음에 대한 차이로 여길 수 있다. 즉, Han et al.(2017)의 연구에서는 자음-모음-자음 형태의 단음절어가 포함된 자극음을 제시한 반면에 본 연구에서는 18개의 한국어 자음과 1개의 동일한 모음 /a/를 결합한 형태의 자극음을 제시했기 때문에 음소 간 분별 혹은 인지에 영향을 미치는 요인에서의 차이라고 추정할 수 있다.

난청은 개인의 의사소통 능력을 저하시키며 나아가 개인의 삶의 질이라는 측면에도 부정적인 영향을 미치기 때문에(Katz & White, 2001) 난청인 대상의 음소 간 유사성, 거리, 그리고 말지각 지도에 대한 적용을 통해 의사소통 능력 평가에 대한 필요성과 중요성이 요구된다(Yoon, 2007). 본 연구에서는 연구 참가자가 모두 20대의 정상 청력을 가진 성인이었으며 이는 음소 간 유사성과 거리 및 말지각 지도 결과에 대한 연령과 난청에 대한 고려를 하지 못했다는 제한점이 있다. 이러한 제한점에도 불구하고, 본 연구의 결과는 정상 청력을 가진 성인을 대상으로 음소 간 분별 및 인지에 대한 결과를 오차행렬과 Shepard's law에 근거하여 구체적으로 분석한 한국어 기반의 최초 연구라는 데 의의를 둘 수 있다. 향후 난청인 대상으로의 심도 있는 후속 연구를 진행하여 난청인의 재활을 도울 수 있는 말지각 지도를 시각적으로 적용할 수 있도록 객관화하고 청능사들이 보다 손쉽게 임상에서 적용할 수 있도록 공학적 접목을 통해 사용

자 친화적 도구로 자리매김하는 것이 중요하다.

중심 단어 : 음소지각 유사성 · 음소지각 거리 · 혼동 음소 ·
소음 속 말지각.

Ethical Statement

All participants signed an informed consent form before conducting the experiments. The protocol of this study had been approved by Institutional Review Board of Hallym University (#HIRB-2014-40; #HIRB-2018-009).

Acknowledgments

N/A

Declaration of Conflicting Interests

There are no conflict interests.

Funding

This work was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2018 S1A3A2074932).

Author Contributions

All authors contributed equally to this work. C.K. designed performed experiments, analyzed data and wrote the paper; S.Y. analyzed data and performed experiments; W.H. designed experiments, analyzed data, and wrote the paper; D.L. and S.M. designed and performed experiments; S.P. and Y.K. designed and analyzed data. Also, the authors discussed the results altogether and implications and commented on the manuscript at each stage.

ORCID iDs

Chanbeom Kwak <https://orcid.org/0000-0001-5657-7536>
Woojae Han <https://orcid.org/0000-0003-1623-9676>

REFERENCES

- An, L. J., Kim, J. S., & Pae, S. Y. (2002). The study on developing a test of speech perception in noise. *Korean Journal of Audiology*, 6(2), 118-127.
- Beattie, R. C. (1989). Word recognition functions for the CID W-22 test in multitalker noise for normally hearing and hearing-impaired subjects. *The Journal of Speech and Hearing Disorders*, 54(1), 20-32.
- Ching, T. Y., Dillon, H., Katsch, R., & Byrne, D. (2001). Maximizing effective audibility in hearing aid fitting. *Ear and Hearing*, 22(3), 212-224.
- Chun, H., Ma, S., Han, W., & Chun, Y. (2015). Error patterns analysis of hearing aid and cochlear implant users as a function of noise. *Journal of Audiology and Otology*, 19(3), 144-153.
- Firszt, J. B., Holden, L. K., Skinner, M. W., Tobey, E. A., Peterson, A., Gagli, W., et al. (2004). Recognition of speech presented at soft to loud levels by adult cochlear implant recipients of three cochlear implant systems. *Ear and Hearing*, 25(4), 375-387.
- Han, W., Chun, H., Kim, G., & Jin, I. K. (2017). Substitution patterns of phoneme errors in hearing aid and cochlear implant users. *Journal of Audiology and Otology*, 21(1), 28-32.
- Johnson, K. (2003). Speech perception. In Johnson, K. *Acoustics and Auditory Phonetics* (pp. 59-78). Maldon, MA/Oxford: Blackwell.
- Katz, J. & White, T. P. (2001) Introduction to the handicap of hearing impairment. In Hull, R. H. (4th ed.). *Aural Rehabilitation: Serving Children and Adults*. San Diego, CA: Singular Thomson Learning.
- Killion, M. C. (1997). Hearing aids: Past, present, future: Moving toward normal conversations in noise. *British Journal of Audiology*, 31(3), 141-148.
- Kim, J. M. (2002). English pronunciation dictionaries for Korean EFL learners. *Jungang Journal of English Language and Literature*, 44(2), 59-91.
- Lee, J. (2016). Standardization of Korean speech audiometry. *Audiology and Speech Research*, 12, S7-S9.
- Lee, J. & Han, J. H. (2019). The effect of noise on cortical auditory evoked potential in cochlear implant users. *Audiology and Speech Research*, 15(1), 38-48.
- Miller, G. A. & Nicely, P. E. (1955). An analysis of perceptual confusions among some English consonants. *The Journal of the Acoustical Society of America*, 27(2), 338-352.
- Plomp, R. (1986). A signal-to-noise ratio model for the speech-reception threshold of the hearing impaired. *Journal of Speech and Hearing Research*, 29(2), 146-154.
- Shepard, R. N. (1972). Psychological representation of speech sounds. In David, E. E. & Denes, P. B. *Human Communication: A Unified View*. (pp. 67-113). New York, NY: McGraw Hill.
- Tye-Murray, N. (2014). *Foundations of Aural Rehabilitation: Children, Adults, and Their Family Members*. (4th ed.). Clifton Park, NY: Delmar Cengage Learning.
- Wilson, R. H. & Strouse, A. (2002). Northwestern University Auditory Test No. 6 in multi-talker babble: A preliminary report. *Journal of Rehabilitation Research and Development*, 39(1), 105-113.
- Yoon, M. S. (2007). A qualitative study of the evaluation of communication abilities in the diagnosis of children with hearing impairment. *Communication Sciences and Disorders*, 12(3), 447-464.