

Comparisons of the Korean Band-Importance and Absolute Transfer Functions according to Speakers' Gender

Sohee Heo¹, Hongyeop Oh¹, In-Ki Jin²

¹Department of Speech Pathology and Audiology, Graduate School, Hallym University, Chuncheon, Korea

²Division of Speech Pathology and Audiology, Research Institute of Audiology and Speech Pathology, College of Natural Sciences, Hallym University, Chuncheon, Korea

화자의 성별에 따른 한국형 주파수중요함수 및 절대전달함수 비교

허소희¹ · 오홍엽¹ · 진인기²

한림대학교 일반대학원 언어병리청각학과¹, 한림대학교 자연과학대학 언어청각학부 · 청각언어연구소²

Purpose: The speech intelligibility index (SII) represents the contribution of audible speech information within given frequency bands to speech intelligibility. Generally, speech intelligibility performance depends on the gender of the speaker. The purpose of this study was to compare how a speaker's gender affects band-importance functions (BIFs) and absolute transfer functions (ATFs), which are important components of the SII. **Methods:** This study consisted of 78 subjects. The SII data for male speakers was obtained from a previous study. The SII data for female speakers was obtained using the same procedure as the previous study. Hearing-In-Noise Test sentences were used as stimuli. Intelligibility scores were measured in 42 filters at 5 signal-to-noise ratios. The BIFs and ATFs were derived using a nonlinear optimization procedure, and then the BIFs and ATFs of the different genders were compared. **Results:** The BIF and ATF of different genders have different characteristics. In the case of the BIF, the gap between the different genders ranged between 0.2% (1,170 Hz) and 4.0% (8,500 Hz) depending on the frequency band. In the case of the ATF, the female speakers' ATF was steeper than the male speakers' ATF. **Conclusion:** The results of the current study indicate that gender can be an important factor for predicting intelligibility. Female speakers' speech seems particularly more intelligible than male speakers' speech when measured at the same audibility levels in Korean SII modeling.

Key Words: Speech intelligibility index, Gender effect, Speech perception, Band-importance function, Absolute transfer function.

Received: August 17, 2018 / **Revised:** October 4, 2018 / **Accepted:** October 17, 2018

Correspondence: In-Ki Jin, Division of Speech Pathology and Audiology, Research Institute of Audiology and Speech Pathology, College of Natural Sciences, Hallym University, 1 Hallymdaehak-gil, Chuncheon 24252, Korea

Tel: +82-33-248-2221 / **Fax:** +82-33-256-3420 / **E-mail:** inkijin@hallym.ac.kr

INTRODUCTION

어음명료지수(speech intelligibility index, SII)는 주어진 주파수 대역의 가청 능력을 산술적으로 계산함으로써 어음인지 수행력을 예측할 수 있는 모델이다[American National Standards Institute (ANSI), 1997/R2012; Amlani et al., 2002]. SII는 주로 보청기 착용 전후의 예측 어음인지 수행력의 변화를

통하여 보청기 적합의 예후를 예측하는 모델로 활용하고 있다. 특히 SII를 청력도상에서 간편하게 눈으로 확인할 수 있도록 제작된 시각화 SII (graphical SII) (Mueller & Killion, 1990)는 주파수별 어음역동범위(dynamic range of speech)와 주파수중요함수(band-importance function, BIF)를 시각적 정보로 제공함으로써 환자 및 보호자가 대상자의 가청 영역 및 어음인지도 수행력에 대해 쉽게 이해하도록 도와주어 상담 시에 효과적인 도구로 사용할 수 있다(Hornsby, 2004; Killion & Mueller, 2010). SII를 산술식으로 표현하면 다음과 같다.

© 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.

$$SII = \sum_{f=1}^F BIF_f \times BAF_f \quad [1]$$

F는 나누어진 주파수 대역의 총 개수를 나타낸다. BIF_f는 해당 주파수 대역의 BIF를 나타내며, BAF_f는 해당 주파수 대역의 대역가청함수(band-auditory function, BAF)를 나타낸다. BIF의 경우, 전체 주파수에 대한 중요도를 1로 설정하고 각 주파수 대역에서 상대적인 중요도를 할당하여 주파수별 중요도를 산정한다. BAF는 주어진 주파수 대역에서 청취자가 청취 가능한 어음에너지의 정도를 수치로 정량화한 값을 의미한다(Amlani et al., 2002; ANSI, 1997/R2012). 예를 들어, 해당 주파수 영역에서 50%만큼 어음 에너지를 들을 수 있다면, 0.5로 계산된다. 총 SII 값은 각 주파수 대역의 BIF와 BAF 수치를 곱한 값들의 총합으로 계산할 수 있다.

SII의 어음인지도 예측은 SII 값과 백분율로 표시된 어음인지도 점수 사이의 관계를 보여주는 절대전달함수(absolute transfer function, ATF)를 이용하여 산출할 수 있다. ATF의 도출 과정을 산술식으로 표현하면 다음과 같다.

$$S = (1 - 10^{-\frac{PA}{Q}})^N \quad [2]$$

S는 어음인지도, P는 청자와 화자의 숙련도, A는 equation [1]에서 계산된 SII 값을 의미한다. Q와 N은 함수의 곡선을 적은 왜곡으로 산출하기 위한 상수(fitting constant) 값을 나타낸다.

SII의 중요한 구성요소인 BIF와 ATF는 다양한 요소들에 의해서 영향을 받는다. 예를 들면, BIF의 경우 측정에 사용되는 언어 및 자극 종류에 따라 주파수별 특성이 다르게 나타난다(DePaolis et al., 1996; Narne et al., 2016; Wong et al., 2007). Narne et al.(2016)과 Wong et al.(2007)은 각각 말라얄람어와 광동어를 사용하여 BIF를 도출하였으며, 영어를 기반으로 개발된 기존 BIF와는 다른 주파수별 중요도 값을 가지고 있다고 보고하였다. 또한 Healy et al.(2013)은 동일 언어(영어)에서 자극음의 종류가 바뀌면 BIF 또한 달라진다고 보고하였다. ATF의 경우에도 사용된 자극음의 종류에 따라 예측 곡선(prediction curve)이 달라지는 것을 확인한 연구들이 있었다(Bell et al., 1992; Depaolis et al., 1996). Depaolis et al.(1996)은 단어, 문장, 담화 등 자극음의 종류에 따라서 ATF가 다르게 도출된다는 것을 확인하였으며, Bell et al.(1992)은 예측 가능한 정도가 다른 두 문장 목록의 ATF 간에 차이가 있음을 보고하였다. 선행연구들의 결과들은 SII의 중요한 구성요소인 BIF 및 ATF는 언어, 자극음의 종류 및 자극음의 예측 가능성 등 다양한 요인에 의해 영향을 받을 수 있음을 시사한다.

본 연구는 BIF와 ATF에 영향을 미칠 수 있는 요소로서 화

자의 성별에 초점을 두었다. 일반적으로 화자의 성별은 어음인지도에 영향을 미치는 중요한 요소로 보고되고 있다(Bradlow et al., 1996; Kim & Lim, 2006). Bradlow et al.(1996)의 보고에 따르면 영어를 모국어로 사용하는 청자와 화자를 대상으로 실험한 결과 여성 화자 10명의 목소리로 검사한 문장인지도가 평균 89.5%로 남성 화자 10명의 목소리로 검사한 문장인지도 평균 점수(86.2%)보다 유의미하게 높은 것을 확인하였다. Kim & Lim(2006)은 남녀 화자의 목소리로 녹음된 한국어 단음절 표를 활용하여 20명의 건청 성인을 대상으로 단음절 어음인지를 측정 한 결과, 남성 화자의 목소리로 측정하였을 때 높은 어음인지를 보이는 단음절과 여성 화자의 목소리로 측정하였을 때 높은 어음인지를 보이는 단음절들이 다름을 보고하였다. 화자의 성별에 따른 어음인지 수행력의 차이는 화자의 성별에 따른 음향적인 특징에서도 예측해 볼 수 있다. Noh & Lee(2012)는 건청 성인 남녀 각 10명의 목소리를 녹음하여 장기평균어음스펙트럼(long-term average speech spectrum)의 값이 화자의 성별에 따라 다르다는 것을 보고하였다. 또한, Klatt & Klatt(1990)는 남녀의 성별에 따라 모음의 기본 주파수(fundamental frequency)가 다르다고 보고하였다. 이러한 선행 연구들의 결과는 화자의 성별 요소가 어음인지도에 영향을 미칠 수 있다는 근거를 뒷받침한다고 볼 수 있다.

SII에서 중요한 구성요소인 BIF와 ATF는 다양한 주파수 필터링 및 신호대잡음비(signal-to-noise ratio) 상황에서 측정된 어음인지도 수행력에 기반하여 도출된다(Amlani et al., 2002; ANSI, 1997/R2012). 만약 화자의 성별에 따라 각 조건에서 측정된 어음인지도 점수가 다르다면, 다양한 상황에서 측정된 어음인지도 점수를 기반으로 도출되는 BIF 및 ATF 값 또한 달라질 수 있다. 따라서 본 연구의 목적은 여성 화자의 목소리로 BIF와 ATF를 구하고 기존의 남성 화자의 데이터와 비교하여 성별에 따른 차이를 확인하고, SII 예측에 미치는 영향을 비교해 보고자 하였다. 만약 성별에 따른 차이가 BIF 및 ATF에서 나타난다면 SII 산출에 성별이 중요한 요소라는 점을 확인할 수 있을 것이며, 차이가 없거나 적어서 어음인지도 예측 값이 유사하게 산출된다면 SII를 산출할 때 화자의 성별은 중요한 고려 대상이 아님을 확인할 수 있을 것이다. 본 연구에서 여성 화자의 목소리 기반 BIF와 SII의 산출은 남성 화자를 대상으로 SII를 산출한 선행연구(Jin et al., 2015)와 동일한 방식에서 이루어졌다. 본 연구의 결과는 화자의 성별에 따른 한국어형 SII의 예측 값의 차이 유무에 대한 정보를 제공할 것이다.

MATERIALS AND METHODS

연구 대상

본 연구의 어음인지도 데이터는 한국어를 모국어로 사용하는 78명(남자 42명, 여자 36명)의 건청 성인을 대상으로 실시하였다(평균연령: 24.3세, 연령분포: 20~29세). 참여자들의 순음 청력검사 결과, 모든 참여자들의 청력역치는 250~8,000 Hz 영역에서 20 dB HL 내의 정상 범위를 보였다. 또한, 고막운동도검사에서 모두 A 타입으로 정상임을 확인하였다. 참여자들에게는 적절한 보상을 지급하였으며, 실험에 참여하기 전에 모든 참여자들은 연구의 목적 및 진행 절차 등에 대한 충분한 설명을 제공받았으며 각 참여자의 동의 후에 실험을 진행하였다. 또한 본 연구의 여성 화자 자극음의 녹음은 음성장애가 없는 건청 성인 여성 35명을 대상으로 진행하였다.

자극음

남성 화자에 의해 산출된 BIF 및 ATF 데이터는 Jin et al.(2015)의 데이터를 사용하였다. Jin et al.(2015)에서 화자로 참여한 한 명의 남성 화자는 다수의 남성 화자에 의해 동일한 문장표를 녹음한 후 녹음에 참여한 남성 화자들의 장기평균어음스펙트럼 값에 가장 근접한 대상자를 화자로 선정하여 연구를 진행하였다. 따라서 본 연구의 자극음은 총 35명의 여성 화자를 대상으로 Korean Hearing In Noise Test (K-HINT) 문장표(Moon et al., 2008)를 녹음한 후 장기평균어음스펙트럼을 산출하여 평균값에 가장 근접한 값을 나타내는 대상자의 목소리로 녹음된 음원을 자극음으로 설정하였다. 자극음의 음성 샘플은 이중벽 방음실(double-wall sound booth)에서 음성분석프로그램(Computerized Speech Lab; PENTAX Medical, Montvale, NJ, USA)과 마이크(Sennheiser e-835s; Sennheiser, Wedemark, Germany)를 사용하여 녹음하였다. 음원의 표준화 주파수를 44,100 Hz로 설정하였으며 양자화는 16 비트(bit)로 설정하였다. 녹음된 음원은 65 dB SPL로 정규화한 후, speech shaped noise를 사용하여 다양한 신호대잡음비(-8, -4, 0, +4, +8 dB)로 제작하였다. 이 다섯 가지 신호대잡음비는 남성 화자의 목소리로 한국어 문장을 이용하여 BIF를 도출한 선행연구에서 인지도 점수와 신호대잡음비 사이에 유의한 효과를 보여주는 효과적인 조건으로 나타났다(Jin et al., 2015). 따라서 본 연구에서도 같은 신호대잡음비를 사용하였다. 또한 음원들은 21개의 저주파수통과필터(200, 300, 400, 510, 630, 770, 920, 1,080, 1,270, 1,480, 1,720, 2,000, 2,320, 2,700, 3,150, 3,700, 4,400, 5,300, 6,400, 7,700, 9,500 Hz)와 21개의 고주파수통과필터(100, 200, 300, 400, 510, 630, 770, 920, 1,080, 1,270, 1,480, 1,720, 2,000, 2,320, 2,700, 3,150, 3,700, 4,400, 5,300,

6,400, 7,700 Hz)에 의하여 필터링되었다. 42개의 필터링 조건은 SII 중요 밴드(SII critical-band) 계산 절차와 동일하게 적용하였다(ANSI, 1997/R2012). 사용된 필터의 기울기(slope)는 옥타브 단위당 96 dB로 설정하였다.

실험 장비

제시음은 청각검사기기(GSI 61; Grason-Stadler, Eden Prairie, MN, USA)로 전달되어 헤드폰(TDH-50P; Telephonics Corporation, Farmingdale, NY, USA)을 통하여 참여자의 오른쪽 귀로 제시되었다. 헤드폰의 출력 레벨은 1 kHz 순음으로 보정하였다.

연구 절차

모든 검사 절차는 Jin et al.(2015)에서 사용한 절차와 동일하게 진행하였다. 이 절차는 Studebaker & Sherbecoe(1991), Wong et al.(2007)의 연구에서 사용한 것과 유사한 실험 프로토콜을 통하여 설계되었다. 참여자들은 헤드폰을 통하여 들은 문장을 들은 만큼 최대한 받아쓰도록 지시받았다. 인지도는 문장의 중심 단어를 정확히 받아쓴 비율로 계산하였다. 문장의 듣기 조건은 저주파수통과필터(low-pass filter, LP filter)와 고주파수통과필터(high-pass filter, HP filter) 각각 21개와 5개의 신호대잡음비를 사용하여 제작된 총 210가지였다. 하지만 본 연구의 실험 설계에 따라 208개의 조건에서만 실험을 진행하였다(Jin et al., 2015). 제외된 두 가지 조건은 -8 dB 신호대잡음비로 제작된 LP 200 Hz 조건과 HP 7,500 Hz 조건이었으며, 모두 0%의 점수가 예상되는 조건이었다. 2가지 실험 조건을 제외한 이유는 자극음(문장) 특성상 학습 효과를 방지하기 위하여 연구 참여자는 동일한 문장에 한 번 이상 노출되지 않도록 실험 디자인을 설계해야 한다. 따라서 대상자당 총 문장표 수(24개 리스트)에 해당하는 24개 조건에만 노출되도록 하기 위해서는 2가지 실험 조건(210 - 2 = 208개 조건)의 제외가 불가피했다(78명 × 24개 리스트 = 208개 조건 × 9개 조건별 데이터 = 1,872개 데이터). 이는 한국 남성 화자의 BIF를 구한 선행연구와 동일한 실험 디자인이었다(Jin et al., 2015). K-HINT는 24개 목록(총 240 문장)으로 구성되어 있기 때문에 각 참여자들은 학습 효과(learning effect)를 방지하기 위하여 총 208개의 조건 중에서 24개의 조건에만 무작위로 참여하였다. 따라서 각 조건당 9개의 데이터가 얻어졌다. 실험 시간은 약 1시간 정도 소요되었으며, 참여자들이 원할 경우 언제든지 휴식을 취할 수 있도록 안내하였다.

통계 분석

피어슨 상관계수를 사용하여 예측된 SII 값과 평균 인지도 점수 사이의 상관관계를 확인하고자 하였다. 통계분석은 SPSS 22

(IBM Corp., Armonk, NY, USA) 소프트웨어를 통해 분석하였고 유의수준은 $\alpha = 0.05$ 이하로 설정하였다.

BIF와 ATF의 도출을 위한 계산방법

BIF 및 ATF의 도출은 남성 화자가 녹음한 K-HINT 문장표로 도출한 결과와 본 연구의 결과를 비교하기 위하여 Jin et al.(2015)에서 사용한 절차를 동일하게 적용하였다. Kates(2013)가 개발한 이 방법은 데이터 평활화(data smoothing) 및 그래픽 설계(graphical construction)를 거치지 않고 BIF와 ATF를 도출하고, 매트랩 최적화 도구(MATLAB version R2013a; Math-Works Inc., Natick, MA, USA)에서 제공하는 비선형 최적화 함수를 기반으로 BIF 및 ATF를 도출하는 방법이다. 최적화의 기준은 equation [2]의 평균제곱근(root mean square) 오차를 최소화하고 검사 조건 전체에서 확인된 어음인지도의 평균 비율을 통합하는 것이었다. 또한 수학적 절차는 BIF와 Q와 N의 값을 동시에 조정해서 SII가 대상자의 인지도 점수에 최적화되도록 제작하였다.

매트랩 프로그램을 사용하여 오류를 최소화하는 방법은 다음과 같은 단계들로 구성하였다. 첫째, 오류 최소화는 매개변수인 Q와 N과 함께 BIF의 초기 추정치를 포함하는 벡터로 설정하였다. 둘째, 매트랩 함수인 'fminsearch'를 통하여 초기 추정치보다 오류를 줄인 BIF를 도출하였다. 셋째, 매트랩 함수인 'fmincon'을 통하여 한 번 더 BIF의 오류 최소화 과정을 거쳤다. 이 방법은 Lagarias et al.(1998)에서 함수 값의 오류 최소화를 위해 적용되었던 방법이었다. 내점(interior point)을 찾는 설정은 BIF의 합이 1이어야 하며 모든 매개변수와 각 주파수 대역별 중요도가 0 이상이어야 한다는 것이었다(Byrd et al., 1999). 마지막으로 5점 평활화(five-point smoother)를 사용하여 BIF 결과의 이상치(outlier)를 제거하였다(Marchand & Marmet, 1983). 본 방법은 여러 선행연구(Jin et al., 2015; Kates, 2013;

Narne et al., 2016)에서 신뢰성이 입증되어 본 연구의 BIF와 ATF 도출 절차로 적용하였다.

RESULTS

화자의 성별에 따른 K-HINT 문장표의 BIF 비교는 Table 1에 제시하였다. 두 BIF는 두 주파수 대역을 제외한 대부분의 주파수 대역에서 중요도의 차이는 2% 미만이었다. 2% 이상 차이를 보인 중심 주파수는 450 Hz와 8,500 Hz로 그 차이는 각각 3.6%와 4.0%였다. 여성 화자의 경우 중심 주파수 250 Hz에서 가장 높은 중요도(8.9%)를 보였으며, 남성 화자의 경우 중심 주파수 350 Hz와 450 Hz에서 가장 높은 중요도(8.3%)를 보였다. 두 BIF의 교차 주파수(crossover frequency)는 중심 주파수 기준 1,750 Hz였다.

화자의 성별에 따른 K-HINT 문장표의 ATF는 Figure 1에

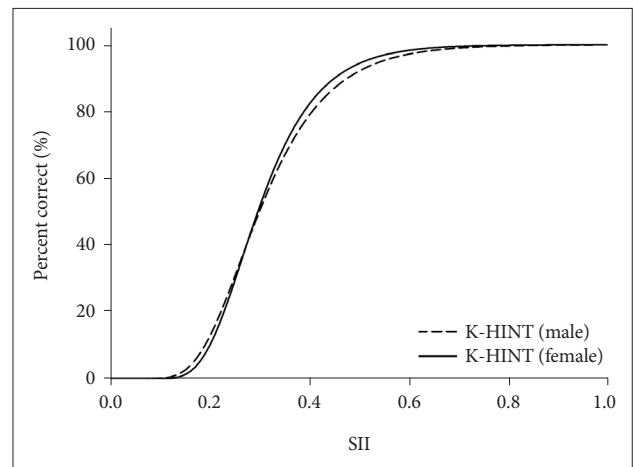


Figure 1. Absolute transfer functions for K-HINT recorded by a male speaker (dashed line) and a female speaker (solid line). K-HINT: Korean Hearing In Noise Test. SII: speech intelligibility index.

Table 1. The BIFs derived from male and female speakers as a function of 21 frequency bands

Band limit (Hz)	CF (Hz)	BIF for female (%)	BIF for male (%)	Band limit (Hz)	CF (Hz)	BIF for female (%)	BIF for male (%)
100–200	150	7.3	5.4	1,720–2,000	1,850	2.3	3.6
200–300	250	8.9	7.4	2,000–2,320	2,150	2.0	4.0
300–400	350	6.2	8.3	2,320–2,700	2,500	3.2	4.7
400–510	450	4.7	8.3	2,700–3,150	2,900	4.8	4.4
510–630	570	5.7	6.7	3,150–3,700	3,400	5.0	3.4
630–770	700	5.3	4.4	3,700–4,400	4,000	4.2	3.5
770–920	840	4.3	2.7	4,400–5,300	4,800	3.8	5.5
920–1,080	1,000	4.4	2.9	5,300–6,400	5,800	4.5	5.9
1,080–1,270	1,170	4.9	4.7	6,400–7,700	7,000	5.2	3.5
1,270–1,480	1,370	4.8	5.4	7,700–9,500	8,500	4.9	0.9
1,480–1,720	1,600	3.7	4.4				

CF: center frequency, BIF: band-importance function

제시하였다. 여성 화자에 의해 도출된 ATF 곡선이 남성 화자에 의해 도출된 ATF 곡선보다 더 가파른 것을 볼 수 있다. 예를 들어, 0.3 SII는 남성 화자의 경우 50.5% 인지도 점수에 대응하는 반면, 여성 화자의 경우 51.8% 인지도 점수에 대응하였다. 또한, 0.45 SII일 때 남성 화자의 경우는 87.0% 인지도 점수에 대응하는 반면, 여성 화자의 경우는 89.9%에 대응하였다.

본 연구에서 도출한 여성 화자의 SII 값과 평균 인지도 점수 간의 산포도는 Figure 2에 제시하였다. 측정된 피어슨 상관관계수는 $r = 0.987$ 로 SII를 통하여 예측한 점수와 실제 측정된 어음인지도 점수 간의 상관관계는 매우 높은 것으로 나타났다. 따라서 본 실험에서 도출된 SII의 예측 정확도가 높다고 볼 수 있다.

DISCUSSIONS

본 연구는 남녀 화자에 의해 녹음된 K-HINT 문장표를 사용하여 도출한 BIF와 ATF를 통하여 성별에 따른 어음인지도 예측 값의 변화 유무를 확인하고자 하였다. 실험 결과 도출된 성별에 따른 BIF는 주파수에 따라 약 0.2%에서 4.0%의 차이를 보였으며, ATF는 여성 화자의 목소리로 도출된 곡선이 남성 화자의 목소리로 도출된 곡선보다 더 가파른 것을 확인할 수 있었다. 곡선의 기울기에 따라 동일한 SII에서 예측되는 어음인지 수행력이 차이가 나타나게 되며, 기울기가 가파를수록 더 높은 수행력이 예측된다. 따라서, 여성 화자의 목소리로 녹음된 K-HINT 문장표가 어음인지 수행에 더 유리한 것으로 볼 수 있다.

영어를 기반으로 성별에 따른 어음인지도 차이를 비교한 연구들은 일반적으로 여성 화자의 목소리가 어음인지에 더 유리

하다는 결과를 보고하였다(Bradlow et al., 1996; Byrd, 1994). Bradlow et al.(1996)은 여성 화자에 의한 평균 어음인지 점수(89.5%)가 남성 화자에 의한 평균 어음인지 점수(86.2%)보다 더 높다는 결과를 보였으며, 이는 본 연구 결과에서 ATF의 0.45 SII와 대응하는 각 어음인지 점수와 비슷한 값(여성 화자: 89.9%, 남성 화자: 87%)이었다. 이러한 결과들은 한국어를 기반으로 실험한 본 연구의 결과와 유사하다고 볼 수 있다.

성별에 따른 BIF 및 ATF 값의 차이가 실제 어음인지도 예측에 얼마나 영향을 주는지 알아보기 위하여 시각화된 SII에 적용하여 확인해 보고자 하였다(Killion & Mueller, 2010; Mueller & Killion, 1990). 시각화된 SII는 청력도상에 가청 영역을 100개의 점으로 표현한다. 한 개의 점은 전체 주파수 범위 중 중요도의 1%(0.01 SII)를 의미한다. 주어진 역동범위(dynamic range) 내에서 주파수 범위별로 계산된 주파수 중요도를 일정한 간격으로 점을 배치시키며 이때, 주파수 중요도를 각각 정수로 표현하고 총합은 100이 되도록 설정해야 한다. 역동범위는 사용된 자극음의 평균제곱근 평균값에서 -15 dB 지점을 최솟값으로,

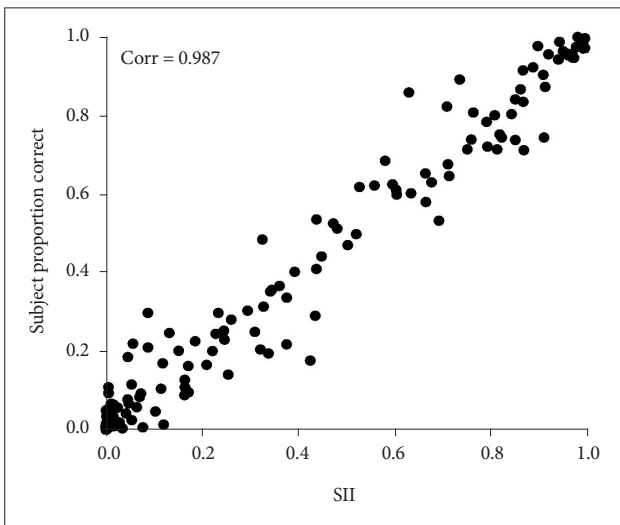


Figure 2. Scatter plot showing the subjects' proportions of correct scores as a function of predicted SII values. The Pearson Corr between the predicted and measured intelligibility scores are displayed in the upper left corner of figure. SII: speech intelligibility index, Corr: correlation coefficient.

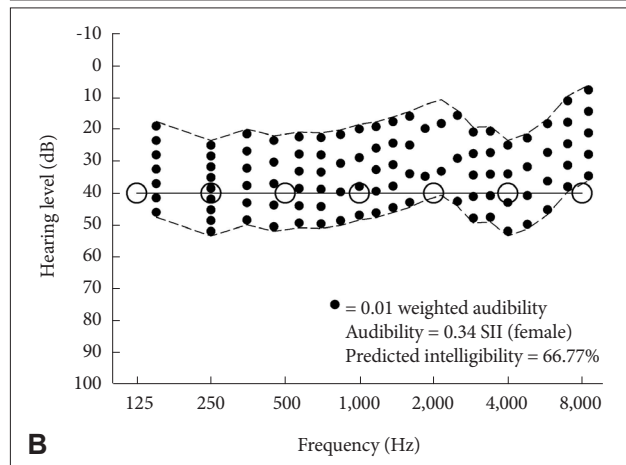
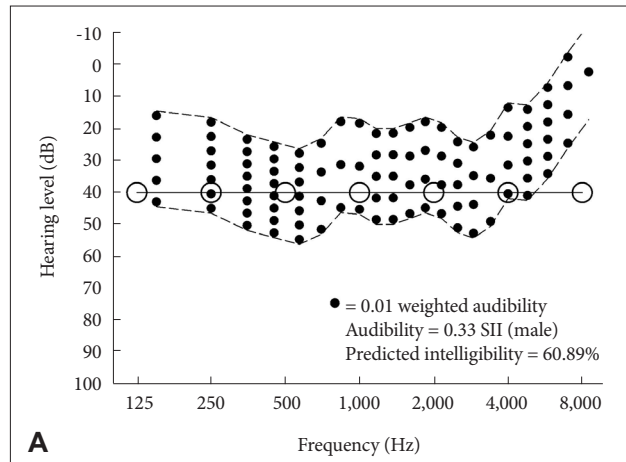


Figure 3. Graphical SII for Korean Hearing In Noise Test recorded by a male speaker (A) and a female speaker (B). SII: speech intelligibility index.

Table 2. Predicted SII values and speech intelligibility score by count-the-dot audiograms for K-HINT

Configuration	HTL (dB) at each frequency band (Hz)								Predicted SII values		Predicted speech intelligibility (%)	
	125	250	500	1,000	2,000	4,000	8,000	F	M	F	M	
Generally slope	20	20	30	40	50	60	70	0.38	0.38	78.0	74.7	
Ski slope	20	30	40	50	60	70	80	0.24	0.23	25.2	23.2	
Rising	60	60	50	40	30	20	10	0.45	0.43	89.9	84.2	

SII: speech intelligibility index, K-HINT: Korean Hearing In Noise Test, HTL: hearing threshold level, F: female, M: male

+15 dB 지점을 최댓값으로 설정하였다(ANSI 1997/R2012). 예를 들어 본 연구 결과에서 여성 화자의 경우를 살펴보면 중심 주파수 250 Hz에서의 주파수 중요도가 8.9%였으므로 9%로 조정하고 해당 범위에 9개의 점을 점선으로 나타낸 역동범위 내에서 일정한 간격으로 나타내었다. 각 주파수별 청력역치를 표시하고 이를 실선으로 이어 실선보다 높은 음압에 위치한 점들을 모두 세면 해당하는 SII를 예측할 수 있다. 한 가지 난청 타입(수평형 난청)에 의한 가청 영역 및 어음인지도 예측 결과는 Figure 3에 제시하였고, 세 가지 난청 타입에 의한 가청 영역 및 어음인지도 예측 결과는 Table 2에 제시하였다. Figure 3에 제시된 자료를 보면 남성 화자와 여성 화자에 의해 예측된 SII 값은 각각 0.33과 0.34로 여성 화자에 의해 예측된 값이 더 높은 것을 확인할 수 있다. ATF (Figure 1)를 통해 예측된 어음인지점수 역시 각각 60.89%와 66.77%로 여성 화자에 의해 예측된 점수가 약 6% 높은 결과를 보였다. Table 2에 제시된 자료 또한, 동일한 역치를 화자 성별에 따라 산출된 시각화 SII에 각각 적용하였을 때 여성 화자에 의해 산출된 예측 SII 값이 더 높거나, 동일한 예측 SII 값이라도 ATF를 통한 예측 어음인지도 점수가 더 높은 것을 확인할 수 있다. 이는 동일한 난청을 가지고 있더라도, 화자의 성별에 따라 어음인지도 수행력이 다를 수 있다는 것을 보여준다. 따라서 본 연구의 결과는 화자의 성별이 SII에서 중요한 요소임을 시사한다.

본 연구에서 사용한 여성 화자 음원은 다화자로 구성된 음원이 아니라 다화자에 의해 녹음된 음원 중 장기평균어음스펙트럼이 평균에 가까운 대상자의 목소리로 녹음된 음원을 사용하여 연구를 진행하였다. 따라서 본 연구에 사용된 한 명의 여성 화자 음원이 평균 여성 장기평균어음스펙트럼을 대변할 수는 있으나 그 외의 음성적 특성이 평균 한국 여성을 대변한다고 보기엔 한계가 있을 수 있다. 하지만 단화자로 구성된 음원을 적용한 데는 몇 가지 이유가 있었다. 첫째, 남성 화자에 의해 도출된 선행연구(Jin et al., 2015)도 본 연구에서 화자를 선정할 동일한 기준으로 연구를 진행하여 정확한 비교 조건을 맞추기 위하여 단화자 음원을 사용하였다. 둘째, 일반적인 SII 연구에서는 대부분 단화자 음원을 사용하여 언어 및 자극음에 따른 비교를 실시하였다(DePaolis et al., 1996; Healy et al., 2013; Wong et al., 2007). 일반적으로 남성과 여성 간의 음향적인 차이가 분명하기

때문에(Amlani et al., 2002), 단화자일지라도 성별에 따른 음향적 특성이 방대한 수(1,872개)의 어음인지도 데이터에 기반하여 산출된 BIF 및 ATF에서 나타날 가능성이 높기 때문이다. 셋째, 장기평균어음스펙트럼은 음성적 특징을 나타내는 가장 보편적인 척도이다(Noh & Lee, 2012). 특히 BIF는 주파수별 중요도를 산출하는 함수이기 때문에 주파수별 에너지 분포를 나타내는 여성 장기평균어음스펙트럼의 평균에 해당하는 대상자의 음원은 여성의 보편적인 음성적 특성을 나타낸다고 볼 수 있다. 하지만 더욱 정확한 성별에 따른 차이를 확인하기 위해서는 다화자로 구성된 음원을 기준으로 동일한 실험 조건에서 BIF 및 ATF를 도출하는 후속 연구가 필요할 수 있다.

본 연구는 BIF와 ATF를 통한 SII의 예측이 화자의 성별에 따라 예측 결과에 영향을 미치는지 확인하고자 하였다. 그 결과, 성별에 따른 BIF는 주파수에 따라 적게는 0.2%, 크게는 4.0% 정도 차이를 보였다. 또한, ATF는 여성 화자의 목소리로 도출된 곡선이 조금 더 가파른 것을 확인할 수 있었다. 이러한 결과는 동일한 주파수별 난청일지라도 화자의 성별에 따라 어음인지도가 달라질 수 있는 가능성을 암시한다. 본 연구의 결과가 일반적인 성별에 따른 어음인지도 연구 결과와 유사하다는 측면에서 본 연구의 타당성을 지지할 수 있을 것이다. 본 연구의 결과는 SII를 활용한 보청기 적합 예측의 예측에서 화자의 성별이 중요한 요소가 될 수 있음을 시사한다.

중심 단어 : 어음명료지수 · 성별 효과 · 어음인지 · 주파수중요함수 · 절대전달함수.

Ethical Statement

The study was approved by the Institutional Review Board of Hallym University (HIRB-2017-002).

Acknowledgments

The authors thank Kyungju Lee for recording stimuli. They also thank Dr. James M. Kates and Dr. Kathryn H. Arehart for providing specific information at the step of study design development.

Declaration of Conflicting Interests

There are no conflict interests.

Funding

This study was supported by Basic Science Research Program Grant NRF-2015R1C1A1A01052458 funded by the Ministry of Science, ICT and Future Planning.

REFERENCES

- American National Standards Institute. (1997). *ANSI S3.5-1997 (R2012): Methods for Calculation of the Speech Intelligibility Index*. New York, NY: Acoustical Society of America.
- Amlani, A. M., Punch, J. L., & Ching, T. Y. (2002). Methods and applications of the audibility index in hearing aid selection and fitting. *Trends in Amplification, 6*(3), 81-129.
- Bell, T. S., Dirks, D. D., & Trine, T. D. (1992). Frequency-importance functions for words in high- and low-context sentences. *Journal of Speech, Language, and Hearing Research, 35*(4), 950-959.
- Bradlow, A. R., Torretta, G. M., & Pisoni, D. B. (1996). Intelligibility of normal speech I: Global and fine-grained acoustic-phonetic talker characteristics. *Speech Communication, 20*(3-4), 255-272.
- Byrd, D. (1994). Relations of sex and dialect to reduction. *Speech Communication, 15*(1-2), 39-54.
- Byrd, R. H., Hribar, M. E., & Nocedal, J. (1999). An interior point algorithm for large-scale nonlinear programming. *SIAM Journal on Optimization, 9*(4), 877-900.
- DePaolis, R. A., Janota, C. P., & Frank, T. (1996). Frequency importance functions for words, sentences, and continuous discourse. *Journal of Speech, Language, and Hearing Research, 39*(4), 714-723.
- Healy, E. W., Yoho, S. E., & Apoux, F. (2013). Band importance for sentences and words reexamined. *The Journal of the Acoustical Society of America, 133*(1), 463-473.
- Hornsby, B. W. Y. (2004). The speech intelligibility index: What is it and what's it good for? *The Hearing Journal, 57*(10), 10-17.
- Jin, I. K., Kates, J. M., Lee, K., & Arehart, K. H. (2015). Derivations of the band-importance function: A cross-procedure comparison. *The Journal of the Acoustical Society of America, 138*(2), 938-941.
- Kates, J. M. (2013). Improved estimation of frequency importance functions. *The Journal of the Acoustical Society of America, 134*(5), EL459-EL464.
- Killion, M. C. & Mueller, H. G. (2010). Twenty years later: A new count-the-dots method. *The Hearing Journal, 63*(1), 10, 12-14, 16-17.
- Kim, E. O. & Lim, D. (2006). Effects of word difficulty and talkers on monosyllabic word recognition tests. *Audiology, 2*(2), 102-106.
- Klatt, D. H. & Klatt, L. C. (1990). Analysis, synthesis, and perception of voice quality variations among female and male talkers. *The Journal of the Acoustical Society of America, 87*(2), 820-857.
- Lagarias, J. C., Reeds, J. A., Wright, M. H., & Wright, P. E. (1998). Convergence properties of the Nelder-Mead simplex method in low dimensions. *SIAM Journal on Optimization, 9*(1), 112-147.
- Marchand, P. & Marmet, L. (1983). Binomial smoothing filter: A way to avoid some pitfalls of least-squares polynomial smoothing. *Review of Scientific Instruments, 54*(8), 1034-1041.
- Moon, S. K., Kim, S. H., Mun, H. A., Jung, H. K., Lee, J. H., Choung, Y. H., et al. (2008). The Korean hearing in noise test. *International Journal of Audiology, 47*(6), 375-376.
- Mueller, H. G. & Killion, M. C. (1990). An easy method for calculating the articulation index. *Hearing Journal, 43*(9), 14-17.
- Narne, V. K., Prabhu, P., Thuvassery, P., Ramachandran, R., Kumar, A., Raveendran, R., et al. (2016). Frequency importance function for monosyllables in Malayalam. *Hearing, Balance and Communication, 14*(4), 201-206.
- Noh, H. & Lee, D. H. (2012). Cross-language identification of long-term average speech spectra in Korean and English: Toward a better understanding of the quantitative difference between two languages. *Ear and Hearing, 33*(3), 441-443.
- Studebaker, G. A. & Sherbecoe, R. L. (1991). Frequency-importance and transfer functions for recorded CID W-22 word lists. *Journal of Speech, Language, and Hearing Research, 34*(2), 427-438.
- Wong, L. L., Ho, A. H., Chua, E. W., & Soli, S. D. (2007). Development of the Cantonese speech intelligibility index. *The Journal of the Acoustical Society of America, 121*(4), 2350-2361.