

A Comparison of Band-Importance Function and Transfer Function Using Clear Speech and Conversational Speech

Hongyeop Oh¹, Sohee Heo¹, 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

명료어음과 일반 발성법에 의해 산출된 주파수중요함수와 전달함수 비교

오홍엽¹ · 허소희¹ · 진인기²

한림대학교 일반대학원 언어병리청각학과¹, 한림대학교 자연과학대학 언어청각학부 · 청각언어연구소²

Purpose: This study compared band-importance function (BIF) and transfer function (TF) derived using clear and conversational speech. **Methods:** There were 156 participants in this study: 78 for the clear speech group and the others for the conversational speech group. The data used to derive BIF and TF were obtained using the same procedure as the previous study. Sentences from Speech Perception In Noise test were used as stimuli. Intelligibility scores were measured via 42 filters at five signal-to-noise ratios. The BIFs and TFs for clear and conversational speech were derived using a nonlinear optimization procedure. Then those for the different production methods were compared. **Results:** The BIF and TF for different production methods have different characteristics. For BIF, the highest weights of importance for clear speech ranged between 200 Hz and 400 Hz and the highest weights for conversational speech ranged between 510 Hz and 770 Hz. The maximum different band-importance weight for the two production methods was 1.92%p at a center frequency of 2,150 Hz. The TF for clear speech was steeper than the TF for conversational speech. **Conclusion:** The results of the current study provide information on the degree to which the clear speech is more intelligible according to the listener's audibility compared to that for conversational speech. Audiologists and hearing professionals will use this information, when considering the communication strategy for people with hearing loss.

Key Words: Speech intelligibility index, Clear speech, Band-importance function, Transfer function, Speech intelligibility.

Received: May 28, 2019 / **Revised:** July 17, 2019 / **Accepted:** July 18, 2019

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

명료어음(clear speech)은 대화 상황에서 말소리 이해도를 높일 수 있는 발성법으로써 일반 발성법과 비교하였을 때, 말소리 산출 속도(speech rate)의 감소, 음의 높낮이(pitch) 범위의 확장, 중요 어휘의 강도(intensity) 증가와 같은 음향학적 변조 특성을 가진 조음 수정 전략으로 설명할 수 있다(Smiljanić &

Bradlow, 2007). 명료어음 사용에 따른 가장 큰 효과는 어음인지도 향상이며 연령, 청력, 언어와 무관하게 청취자의 어음인지도를 향상시킬 수 있는 화자 중심(talker-centered) 의사소통 중 하나로 볼 수 있다. Uchanski(2005)의 연구에 따르면 5 dB의 신호대잡음비(signal-to-noise ratio, SNR) 상황에서 건청 성인에게 명료어음을 사용하여 어음인지도를 측정할 결과, 일반 발성법에 비해서 15%에서 20% 정도의 어음인지도 향상이 나타난다고 보고하였다. Bradlow et al.(2003)의 연구에 따르면 소음 환경에서 학습장애가 있는 아동과 건청 아동에게 명료어음을 사용하여 어음인지도를 측정할 결과, 일반 발성법에 비

© 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.

해서 건청 아동 그룹은 8.8%, 학습장애 아동 그룹은 9.2%의 어음인지도 향상이 나타난다고 보고하였다. 또한 Picheny et al. (1985)의 연구에 따르면 감각신경성 난청인에게 무의미 문장을 사용하여 명료어음으로 어음인지를 측정할 결과, 일반 발성법에 비해서 17% 정도의 어음인지도 향상이 나타난다고 보고하였다. Payton et al.(1994)의 연구에 따르면 건청인과 난청인을 대상으로 무의미 문장을 사용하여 무향 환경(anechoic environment), 거실 환경, 회의실 환경 상황에서 명료어음으로 어음인지를 측정하였을 때 일반 발성법에 비해서 건청인 그룹은 약 20%, 난청인 그룹의 경우 약 26% 정도의 어음인지도 향상이 나타났다고 보고하였다. Smiljanic & Bradlow(2007)의 연구에 따르면 5 dB의 신호대잡음비 상황에서 영어가 모국어인 화자와 영어 수행력이 높은 크로아티아인을 대상으로 의미론적 변칙 문장을 사용하여 어음인지를 측정할 경우 일반 발성법은 약 54%, 명료어음은 약 70% 정도의 어음인지도 수행력이 나타났다고 보고하였다. 선행연구들의 결과에 따르면 명료어음은 청자의 난청 및 소음하 상황 등 청취 조건이 열악할 수 있는 상황에서 어음인지를 향상시킬 수 있는 의사소통 전략임을 알 수 있다.

대상자의 청력과 소음 환경은 개인의 가청능력(audibility)에 있어 중대한 영향을 미치는 요소이며, 대상자의 가청능력이 달라지면 어음인지도 또한 달라진다(Amlani et al., 2002). 대표적으로 가청능력(audibility)에 따른 어음인지를 예측할 수 있는 모델로는 어음명료지수(speech intelligibility index, SII)가 있다. 어음명료지수는 주어진 주파수 대역의 가청능력을 계산함으로써 어음인지 수행력을 예측할 수 있는 모델이다[American National Standard Institute (ANSI) S3.5 (R2012)] (ANSI, 1997; Amlani et al., 2002). 어음명료지수를 도출하는 데 있어 중요한 구성 요소로는 주파수중요함수(band-importance function, BIF)와 전달함수(transfer function, TF)가 있다. 주파수중요함수는 음성을 인지하는 것에 있어서 주파수의 중요도를 의미하며, 전체 주파수에 대해 중요도를 1로 설정하고 각 주파수 대역에서 상대적인 중요도를 할당하여 주파수별 중요도를 산정할 수 있다[ANSI S3.5 (R2012)] (ANSI, 1997; Amlani et al., 2002). 전달함수는 어음명료지수 값과 백분율로 표시된 어음인지도 점수 사이의 관계를 반영하며, 이를 통해 어음명료지수 값을 어음인지도 예측 값으로 변환할 수 있다[ANSI S3.5 (R2012)] (ANSI, 1997; Amlani et al., 2002).

Jin et al.(2015) 및 Jin et al.(2017)은 남성 화자에 의해 녹음된 한국어 문장표를 기반으로 어음명료지수를 측정하였다. 연구 결과 주파수중요함수의 경우 630 Hz 이하의 저주파수 영역이 전체 주파수 범위에서 약 38% 정도의 중요도로 나타났으며 (Jin et al., 2015), 전달함수의 경우 0.5 SII가 약 92%의 어음인지도로 나타났다(Jin et al., 2017). 또한 Heo et al.(2018)은 여

성 화자에 의해 녹음된 한국어 문장표를 기반으로 어음명료지수를 측정하였다. 연구 결과 주파수중요함수의 경우 630 Hz 이하의 저주파수 영역이 전체 주파수 범위에서 약 33% 정도의 중요도를 보였으며, 전달함수의 경우 0.5 SII가 약 94%의 어음인지도로 나타났다. 한국어를 바탕으로 어음명료지수를 산출한 위 선행연구들은 보통의 발성 노력 및 발성 속도에 의한 일반 발성법으로 녹음된 자극음에 기반하여 연구를 진행하였다. 따라서 일반 발성 상황에서의 어음인지도 예측에는 유용한 모델일 수 있으나, 일반 발성과 비교하여 음향적인 변조(느린 발화 속도, 주파수 스펙트럼 변화 등)가 발생하는 명료어음 상황에서 어음인지를 예측하기에는 한계가 있다. Lee et al.(2018)은 35명의 연구 참가자를 대상으로 동일한 문장표를 일반 발성과 명료어음으로 녹음하여 어음의 역동 범위를 비교하였다. 그 결과 명료어음 역동 범위의 최대어음레벨(maximum level)이 일반 발성법에 의한 레벨에 비해 0.47 dB (570 Hz)에서 2.96 dB (1,170 Hz) 정도 높은 것으로 나타났다. 어음의 역동 범위에서 최대어음레벨의 차이는 가청 영역의 차이를 의미하며, 가청 영역은 어음명료지수의 핵심 구성 요소인 주파수중요함수 및 전달함수의 도출에 직접적인 영향을 미치는 변수이므로 어음명료지수 결과 변화에 직접적인 원인으로 작용한다[ANSI S3.5 (R2012)] (ANSI, 1997). 따라서 명료어음에 의한 어음인지도 효과를 대상자의 가청능력에 따라 예측하기 위해서는 명료어음 기반의 어음명료지수 도출이 요구된다.

본 연구는 일반 발성과 명료어음 기반의 주파수중요함수 및 전달함수 비교를 통하여 발성법에 따른 어음명료지수 예측 값의 영향을 확인하고자 하였다. 선행연구들의 결과를 바탕으로 예상되는 결과는 명료어음에 의한 주파수중요함수의 경우, 일반 발성에 의한 주파수중요함수와 다른 주파수별 중요도를 보일 것으로 예상하며 명료어음 전달함수의 기울기는 일반 발성법에 의한 전달함수에 비해 더 가파른 기울기로 산출될 것으로 예상된다. 따라서 본 연구 결과를 통하여 동일한 가청 영역에서 명료어음이 일반 발성에 비해 어느 정도의 어음인지도 향상 효과가 나타날지 예상할 수 있는 예측 모델을 제시할 수 있을 것이다. 명료어음 기반의 어음명료지수는 난청자의 의사소통 전략 수립에 있어 명료어음이 중재 전략으로 적절한지에 대한 실제적인 가이드라인을 제공할 수 있을 것이다.

MATERIALS AND METHODS

연구 대상

본 연구는 한국어를 모국어로 사용하는 156명의 건청인을 대상으로 두 그룹(명료어음 및 일반 발성법 그룹)으로 나누어 진행하였다. 각 그룹에는 78명씩 본인의 의사와 상관없이 무작

위로 연구 대상자를 배정하였다. 명료어음 그룹의 경우 27명의 남성 청자(평균연령: 23.3세, 연령 분포: 19~27세)와 51명의 여성 청자(평균연령: 21.6세, 연령 분포: 19~25세)가 참여하였다. 일반 발성법 그룹의 경우 42명의 남성 청자(평균연령: 22.7세, 연령 분포: 18~27세)와 36명의 여성 청자(평균연령: 21.8세, 연령 분포 19~25세)가 참여하였다. 모든 연구 대상자는 과거에 이과적 병력이 없었으며, GSI-61 (Grason-Statler Inc., Eden Prairie, MN, USA)과 TDH-50 (Telephonics, Farmingdale, NY, USA) 헤드폰을 사용하여 방음실에서 순음청력검사(pure tone audiometry)를 실시한 결과, 250~8,000 Hz의 주파수 대역에서 20 dB HL 이하의 역치를 보였다. 또한 모든 연구 대상자의 고막운동성검사(tympanometry) 결과는 A형으로 정상적인 고막운동도를 보였다. 본 연구에 참여하기 전에 연구 대상자들에게 연구의 목적, 방법 및 진행 절차에 대한 설명을 제공하였으며 참여자들의 동의 후에 연구를 진행하였다. 또한 본 연구는 음성 녹음 파일 청취를 통해 들은 대로 종이에 작성하는 방식이 주요한 방법이어서 연구 대상자가 연구 진행 중에 휴식을 원하거나 피로를 느낄 때 충분한 휴식 시간을 제공하였다.

자극음

본 연구의 자극음은 음성장애가 없는 35명의 여성 화자를 통해 녹음을 진행하고 장기평균어음 스펙트럼을 산출하여 평균값에 가장 가까운 값을 나타내는 대상자의 목소리를 선정하여 자극음으로 설정하였다. 음성장애의 유무는 Multi-Dimensional Voice Program [Computerized Speech Lab (CSL), Kay-PENTAX™, Montvale, NJ, USA]을 사용하여 정상 범위의 기본 주파수, 음조 섭동(pitch perturbation, Jitter) 및 진폭 섭동(amplitude perturbation, Shimmer)을 기준으로 결정하였다. 자극음은 두 가지(명료어음, 일반 발성) 발성법으로 녹음하였고, 명료어음 발성법은 다양한 명료어음의 발성법 중 한 가지를 선택하였다. 선택된 발성법으로는 한국어에 대해 익숙하지 않은 외국인에게 한국어로 이야기하듯이 발성하는 방법이었으며, 녹음에 참여하는 화자는 본 발성 방법을 따르되 최대한 자연스러운 발화를 하도록 안내를 하였다(Godoy et al., 2014; Picheny et al., 1985; Smiljanić & Bradlow, 2007). 녹음에 사용된 문장으로는 표준화된 한국형 소음하 문장인지(Korean Speech Perception In Noise, K-SPIN) 문장표의 문장 240개를 사용하였다(An et al., 2002). 음성 녹음 시에는 이중벽 방음실(double-wall sound booth) 내에 설치된 음성 분석 프로그램(CSL, KayPENTAX™)과 마이크(Sennheiser e-835s, Sennheiser, Wedemark, Germany)를 사용하여 의자에 편안한 자세로 앉은 상태에서 입은 마이크로부터 10 cm 거리를 유지하여 45도 각도로 녹음을 진행하였다. 녹음에 참여하는 화자는 한국어문

장검사 목록을 자연스럽게 읽기 위한 숙지 과정을 거쳤고, 녹음 시 화자의 발음이나 조음이 부자연스러운 경우를 방지하기 위해 두 명의 청능사가 모니터링에 참여하였다. 두 명의 청능사 중 단 한 명이라도 화자가 말을 더듬었다고 생각하거나 말의 억양과 강세 등을 고려하였을 때, 화자의 발화가 부자연스럽다고(과조음, over-articulation) 느낀 경우 재녹음을 실시하였다. 녹음된 음원은 Adobe Audition 소프트웨어 버전 3.0 (Adobe Systems, San Jose, CA, USA)을 사용하여 표본화 주파수와 비트(bit) 수를 44,100 Hz와 16 비트로 설정하였다. 어음명료지수의 산출에는 다양한 소음 상황에서 측정된 어음인지도 수행력 점수가 요구된다. 따라서 다양한 신호대잡음비 조건을 설정하기 위하여 유사 선행 논문(Heo et al., 2018; Jin et al., 2015)에서 사용된 절차를 참고하였다. 녹음된 음원을 65 dB sound pressure level로 조정하였고 어음형태잡음(speech shaped noise)을 사용하여 다양한 신호대잡음비(-8, -4, 0, +4, +8 dB SNR)로 음원을 제작하였다. 본 연구에서 적용한 신호대잡음비는 남성 화자의 목소리로 녹음된 한국어 문장을 이용하여 주파수중요함수를 도출한 선행연구에서 신호대잡음비와 인지도 사이에 유의한 효과를 보인 조건으로 설정하였다(Jin et al., 2015). 또한 음원은 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)를 사용하여 필터링하였다. 42개의 필터링 조건은 어음명료지수 임계 대역(SII critical-band) 계산 절차를 따랐다[ANSI S3.5 (R2012)] (ANSI, 1997). 모든 필터의 기울기는 옥타브 단위당 96 dB로 설정하였다.

연구 절차

연구 참여자들은 음성의 제곱평균제곱근(root-mean square) 수준이 동일한 1,000 Hz 순음으로 보정된 TDH-50을 통해 제시된 자극음 오른쪽 귀로 청취하였고, 모든 절차는 ANSI S3.1-1999 (ANSI, 2013)의 주변 소음 수준 기준을 만족하는 이중벽 방음실에서 진행하였다. 자극음은 노트북 컴퓨터에서 GSI-61 청력검사기를 통해 제시하였다. 검사의 절차는 연구 참여자가 자극음을 집중하여 듣고 제공되는 종이에 받아쓰는 방식으로 진행하였다. 연구 참여자에게 제시되는 문장의 일부만 들었을지라도 추측하여 가능한 부분까지 받아쓰도록 권고하였다. 인지도는 문장의 키워드(단어)를 정확히 맞춘 수를 기준으로 백분율로 기록하였다. 문장의 듣기 조건은 총 210가지(21개의 저주파수통과필터와 21개의 고주파수통과필터 × 5개의 신호대잡음비)였지만 본 연구의 실험 설계 때문에 208개의 조건

으로 연구를 진행하였다. 제외된 두 조건(-8 dB 신호대잡음비로 제작된 저주파수통과필터 중에 200 Hz, 고주파수통과필터 중에 7,500 Hz)은 인지도가 모두 0%로 예상되는 조건이었다. 본 연구에 사용된 실험 설계에 대한 자세한 설명은 이하 본문에 설명하였다. K-SPIN은 24목록(총 240문장)으로 구성되어 있기 때문에 총 208개의 조건 중 24개의 조건에 무작위로 참여하였다. 따라서 참여자들은 한 번 이상 동일한 문장을 듣지 못하였고 각 조건당 9개의 데이터를 수집할 수 있었다. 각 연구 대상자는 할당된 검사 조건에 대한 인지도를 측정하는 데 약 2시간이 소요되었다. 본 연구는 유사 선행연구인 Jin et al.(2015), Studebaker & Sherbecoe(1991), Wong et al.(2007)의 연구에서 사용한 프로토콜과 동일하였다.

실험 설계

본 연구에서 필요로 하는 대상자 수와 총 210개 조건의 수에 대한 문장의 수의 한계로 인해, 본 연구는 집단 간 설계와 집단 내 설계를 포함하는 혼합 설계를 사용하였다. 또한 본 연구에서는 실험 오차의 분산을 줄이고 완비 블록 설계에서 가능한 것보다 필터링과 신호대잡음비 조건을 정확하게 비교하기 위해 균형불완비블록 설계(balanced incomplete block design)를 사용하였다. 완비블록 설계는 모든 실험 조건이 실험의 각 블록에 표현되는 것을 의미하지만, 실험 조건의 수가 블록당 단위 수를 초과하면 균형불완비블록 설계를 사용해야 한다(Hinkelmann & Kempthorne, 2005). 균형불완비블록 설계는 Yates(1936)가 제안하였고 Fleiss(1981), Jung(1961), Yang et al.(2011) 연구에서 사용하였다. 혼합블록 설계를 사용하기 위해서는 수와 자극음 수의 곱셈은 반복 횟수와 총 조건의 수의 곱셈과 같아야 한다[대상자 수(78) × 자극 수(24) = 반복 횟수(9) × 총 조건의 수(208), 78 = 1,872]. 위 조건을 충족시키기 위해, 210개의 조건에서 0% 점수를 가질 것으로 예상되는 신호대잡음비 -8 dB의 저주파수통과필터 200 Hz와 고주파수통과필터 7,500 Hz를 총 조건에서 제외하여 208개의 조건을 본 연구의 혼합블록 설계로 선택하였다.

주파수중요함수와 전달함수의 산출 방식

주파수중요함수와 전달함수를 구하기 위한 어음명료지수의 산술식은 다음과 같다.

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

F는 나누어진 주파수 영역의 총 수를 나타낸다. BIF_f는 해당 주파수 영역의 주파수중요함수를 나타내며, BAF_f는 해당 주파수 영역의 대역가청함수(band-auditory function)를 나타낸다.

주파수중요함수의 경우 전체 주파수에 대한 중요도를 1로 설정하고 각 주파수 대역에서 상대적인 중요도를 할당하여 주파수별 중요도를 산정한다. 대역가청함수(가청영역함수)는 주어진 주파수 대역에서 청취자가 청취 가능한 어음에너지의 정도를 수치로 정량화한 값을 의미한다[ANSI S3.5 (R2012)] (ANSI, 1997). 예를 들어 해당 주파수 영역에서 30%만큼 어음에너지를 들을 수 있다면 0.3으로 계산한다. 총 어음명료지수 값은 각 주파수 대역의 주파수중요함수와 대역가청함수 수치를 곱한 값들의 총합으로 계산할 수 있다. SII의 어음인지도 예측은 어음명료지수 값과 백분율로 표시된 어음인지도 점수 사이의 관계를 보여주는 전달함수를 이용하여 산출할 수 있다. 전달함수를 구하는 산술식은 다음과 같다.

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

S는 어음인지도, P는 청자와 화자의 숙련도(본 연구에서는 숙련도에 문제가 없는 정상 청력인을 대상으로 하여 1로 설정함), A는 equation [1]에서 계산된 어음명료지수 값을 의미한다. Q와 N은 함수의 곡선을 적은 왜곡으로 산출하기 위한 상수(fitting constant) 값을 나타낸다.

위에서 언급한 산술식들에 근거하여 주파수중요함수 및 전달함수는 여러 선행연구(Heo et al., 2018; Jin et al., 2015; Kates, 2013)에서 사용한 비선형최적화 함수(nonlinear optimization function) 방법(Kates, 2013)을 사용하여 도출하였다. 이 방법은 데이터 평활화(data smoothing) 및 그래픽 설계(graphical construction)의 영향을 받지 않은 인지도 데이터(raw data)를 사용하여 주파수중요함수와 전달함수를 도출할 수 있는 방법이다. 비선형최적화 함수는 매트랩(MATLAB version R2013a, MathWorks, Inc., Natick, MA, USA)에서 제공하며, 어음인지도 값의 제곱평균제곱근 오차를 최소화하고 검사 조건 자체에서 도출된 어음인지도의 평균 비율을 통합할 수 있다. 바꿔 말하면 주파수중요함수 및 전달함수 도출에 사용되는 산술적 공식의 상수(Q, N) 값을 조정하여 어음명료지수의 예측 값과 대상자의 실제 어음인지도 점수 간의 상관관계를 최적화하는 방법이다. 예를 들어 최적화된 어음인지도의 평균은 다섯 개의 신호대잡음비(-8, -4, 0, +4, +8 dB SNR)와 42개의 필터(저주파수통과필터 21개 및 고주파수통과필터 21개)를 주파수중요함수와 전달함수를 도출하는 데 사용하였다. 저주파수통과필터를 사용하여 구해진 다섯 가지의 신호대잡음비별 주파수중요함수를 최적화 과정을 통해 하나의 주파수중요함수로 통합하였으며 고주파수통과필터 또한 동일한 과정을 진행하였다. 그 다음 저주파수통과필터의 주파수중요함수와 고주파수통과필터의 주파수중요함수를 최적화 과정을 통해 통합하였다. 전달함

수 또한 위와 같은 과정을 거쳤고 최종적으로는 최적화된 하나의 주파수중요함수와 전달함수를 도출하였다.

MATLAB 프로그램의 오류 최소화는 매개변수인 두 개의 상수(Q, N) 값과 함께 주파수중요함수의 초기 추정치를 포함하는 벡터로 설정하였으며, 역동 범위는 30 dB로 고정하였다. 초기 매개 변수는 Q 값 0.2178과 N 값 15.52로 설정하였고 21개의 주파수중요함수 가중치는 모두 1/21로 초기화하였다. 또한 MATLAB의 'fminsearch' 함수를 사용하여 초기 추정치보다 오류를 줄인 주파수중요함수를 도출하였다(Lagarias et al., 1998). 이후에 다양한 변수의 함수를 제한적으로 최소화하기 위해 MATLAB의 'fmincon' 함수를 사용하여 두 번째 최소화를 거쳤다. 내점(interior point)을 찾는 설정은 주파수중요함수의 합이 1이어야 하며 모든 매개변수와 각 주파수 대역별 중요도가 0 이상하도록 설정하였다(Byrd et al., 1999). 최종적으로는 5점 이항식 평활화(five-point binomial smoother)를 적용하여 주파수중요함수 결과의 이상치(outlier)를 제거하였다(Marchand & Marmet, 1983).

RESULTS

주파수중요함수의 산출

K-SPIN 문장표를 사용한 명료어음과 일반 발성법 사용에

따른 주파수중요함수는 Table 1에 제시하였다. 명료어음의 주파수중요함수는 중심 주파수 250 Hz의 영역에서 가장 높은 중요도(7.94%)를 보였다. 또한 350 Hz의 중심 주파수 영역에서는 두 번째로 높은 중요도(7.51%)를 보였으며, 700 Hz의 중심 주파수 영역에서는 세 번째로 높은 중요도(6.4%)를 보였다. 교차 주파수(cross of frequency)는 1,170 Hz 부근에서 형성되었고 중심 주파수 1,170 Hz 이하는 중요도의 50.02%로 나타났다. 일반 발성법의 주파수중요함수는 중심 주파수 570 Hz의 영역에서 가장 높은 중요도(7.27%)를 보였다. 또한 700 Hz의 중심 주파수 영역에서는 두 번째로 높은 중요도(7.11%)를 보였으며, 250 Hz의 중심 주파수 영역에서는 세 번째로 높은 중요도(6.8%)를 보였다. 교차 주파수는 1,170 Hz 부근에서 형성되었고 중심 주파수 1,170 Hz 이하는 중요도의 50.39%로 나타났다.

전달함수의 산출 및 측정과 예측인지도의 상관성

명료어음과 일반 발성법에 따른 전달함수는 Figure 1에 제시하였다. 명료어음에 의해 도출된 전달함수 곡선은 일반 발성법에 의해 도출된 전달함수 곡선보다 덜 가파른 것으로 나타났다. 예를 들어 명료어음의 0.3 SII 값은 전달함수 곡선의 51.57% 인지도에 해당하는 반면, 일반 발성법의 0.3 SII 값은 전달함수 곡선의 46.2% 인지도에 해당하였다. 또한 명료어음의 0.45 SII 값은 전달함수 곡선의 84.99% 인지도에 해당하는 반면, 일반 발성법

Table 1. BIFs of clear and conversational speech for Korean Speech Perception In Noise as a function of 21 frequency bands

| Band limit (Hz) | CF (Hz) | BIF for clear speech (%) | BIF for conversational speech (%) |
|-----------------|---------|--------------------------|-----------------------------------|
| 100-200 | 150 | 4.48 | 5.66 |
| 200-300 | 250 | 7.94 | 6.8 |
| 300-400 | 350 | 7.51 | 5.9 |
| 400-510 | 450 | 5.41 | 5.79 |
| 510-630 | 570 | 5.67 | 7.27 |
| 630-770 | 700 | 6.4 | 7.11 |
| 770-920 | 840 | 5.03 | 4.57 |
| 920-1,080 | 1,000 | 3.68 | 3.32 |
| 1,080-1,270 | 1,170 | 3.9 | 3.97 |
| 1,270-1,480 | 1,370 | 4.31 | 4.13 |
| 1,480-1,720 | 1,600 | 4.41 | 3.96 |
| 1,720-2,000 | 1,850 | 4.78 | 3.91 |
| 2,000-2,320 | 2,150 | 5.18 | 3.26 |
| 2,320-2,700 | 2,500 | 4.76 | 3.2 |
| 2,700-3,150 | 2,900 | 4.01 | 4.79 |
| 3,150-3,700 | 3,400 | 4.43 | 5.92 |
| 3,700-4,400 | 4,000 | 5.08 | 5.08 |
| 4,400-5,300 | 4,800 | 4.07 | 4.11 |
| 5,300-6,400 | 5,800 | 2.71 | 4.15 |
| 6,400-7,700 | 7,000 | 2.93 | 4.05 |
| 7,700-9,500 | 8,500 | 3.31 | 3.05 |
| Total | | 100 | 100 |

BIF: band-importance function, CF: center frequency

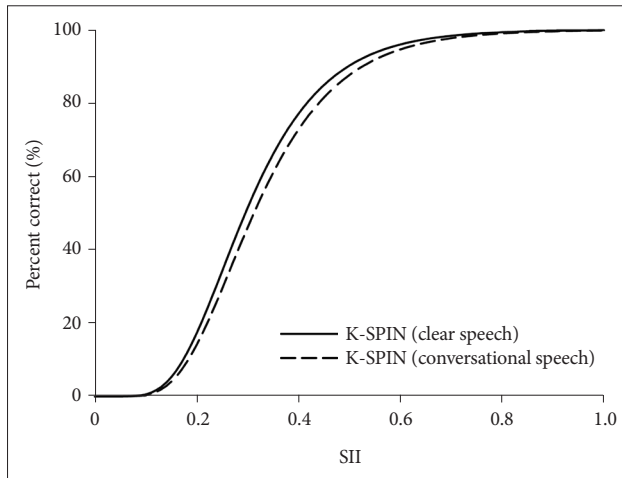


Figure 1. Transfer functions for K-SPIN recorded by clear and conversational speech. X axis represents SII value and Y axis represents keyword correct in percent. K-SPIN: Korean Speech Perception In Noise, SII: speech intelligibility index.

Table 2. Best-fit values for fitting constants of Q, N, Corr, and RMS error for the transfer function of Korean Speech Perception In Noise

| Stimuli | Q | N | Corr | RMS error |
|-----------------------|------|-------|-------|-----------|
| Clear speech | 0.25 | 10.12 | 0.986 | 0.057 |
| Conversational speech | 0.26 | 10.16 | 0.990 | 0.041 |

Q and N : constant. Corr: pearson correlation coefficient, RMS: root mean square

의 0.45 SII 값은 전달함수 곡선의 81.61% 인지도에 해당하였다.

본 연구에서 사용한 피팅 상수 값, 피어슨 상관계수 값 및 제곱평균제곱근 오차의 최적 값은 Table 2에 제시하였다. 높은 피어슨 상관계수 값은 어음명료지수를 통해 예측한 점수와 측정된 어음인지도 점수 간의 상관관계가 높음을 의미하고, 이는 본 연구에서 도출된 명료어음과 일반 발성법 사용에 따른 어음명료지수가 높은 예측 정확도를 보인다는 것을 의미한다. 명료어음의 경우 Q, N 값은 0.25, 10.12로 나타났고, 실제 측정된 어음인지도 점수와 예측된 어음인지도 점수 간의 상관관계는 0.986, 오차는 0.057로 나타났다. 일반 발성법의 경우 Q, N 값은 0.26, 10.16으로 나타났고, 실제 측정된 어음인지도 점수와 예측된 어음인지도 점수 간의 상관관계는 0.990, 오차는 0.041로 나타났다.

DISCUSSIONS

본 연구에서는 K-SPIN 문장표를 사용하여 발성법(명료어음 및 일반 발성법)에 따른 주파수중요함수 및 전달함수를 도출하여 비교하였다. 명료어음으로 산출된 주파수중요함수의 경우 250 Hz (7.94%)에서 350 Hz (7.51%) 사이에서 가장 높은 중요도를 보였으며 교차 주파수는 1,170 Hz 부근이었다. 일반 발성

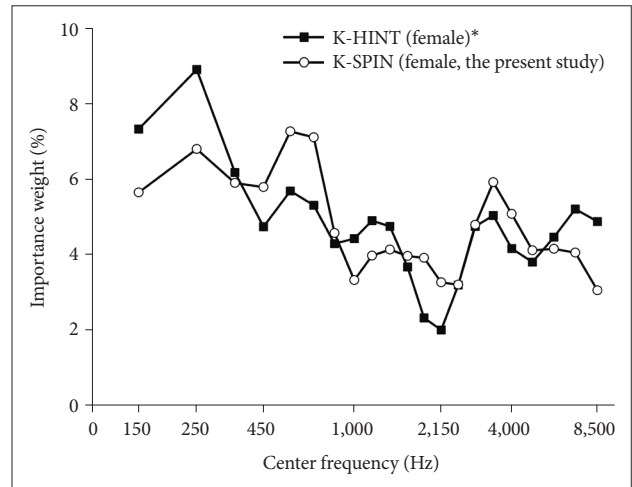


Figure 2. Band-importance functions for K-SPIN (squares) and K-HINT (circles) as a function of 21 frequency bands. *Adapted from Heo et al. Audiology and Speech Research 2018;14:236-242. K-SPIN: Korean Speech Perception In Noise, K-HINT: Korean Hearing In Noise Test.

법으로 산출된 주파수중요함수의 경우 570 Hz (7.27%)에서 700 Hz (7.11%) 사이에서 가장 높은 중요도를 보였으며 교차 주파수는 1,170 Hz 부근이었다. 전달함수의 경우 명료어음으로 산출된 전달함수가 일반 발성법에 의해 산출된 전달함수보다 가파른 기울기를 보였다.

한국어를 기반으로 주파수중요함수 및 전달함수를 산출한 선행연구들은 주로 Korean Hearing In Noise Test (K-HINT) 문장표를 자극음으로 사용하였다(Heo et al., 2018; Jin et al., 2015, 2017). 하지만 본 연구는 문장의 구성 형태가 K-HINT 문장표와는 다른 K-SPIN 문장표를 사용하여 주파수중요함수 및 전달함수를 도출하였다. 특히 Heo et al.(2018)은 본 연구와 마찬가지로 여성 화자의 목소리로 녹음된 자극음을 사용하였고, 본 연구와 동일한 절차에 의해서 주파수중요함수 및 전달함수를 산출하였다. 따라서 자극음의 종류에 따른 차이를 확인하기 위하여 Heo et al.(2018)과 본 연구에서 산출한 주파수중요함수 및 전달함수를 Figures 2와 3을 통해 비교하였다. 주파수중요함수의 경우 K-HINT 문장표는 150 Hz에서 350 Hz 영역이 가장 중요한 주파수 영역이었고, K-SPIN 문장표는 570 Hz에서 700 Hz 영역이 가장 중요한 주파수 영역이었다(Figure 2). 두 주파수중요함수 값 중 가장 큰 차이가 발생한 주파수 영역은 250 Hz로 약 2.1%p (percent point) 차이였으며, 가장 적은 차이가 발생한 주파수 영역은 2,500 Hz로 약 0.01%p 차이였다. 주파수별로는 0.01%p에서 2.1%p 정도 차이가 나타났지만 1,000 Hz에서 2,000 Hz 사이의 중주파수 대역의 중요도가 가장 높았던 영어 문장의 주파수중요함수와 비교하여 저주파수 영역이 더 중요하게 나타나는 공통점도 확인할 수 있다(Healy et al., 2013). 전달함수의 경우 K-HINT 문장표가 K-SPIN에

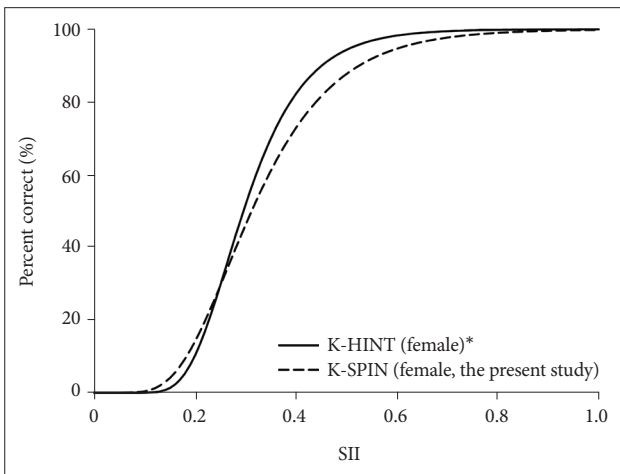


Figure 3. Transfer functions for K-HINT (straight line) and K-SPIN (dotted line). *Adapted from Heo et al. *Audiology and Speech Research* 2018;14:236-242. K-HINT: Korean Hearing In Noise Test, K-SPIN: Korean Speech Perception In Noise, SII: speech intelligibility index.

비해 더 가파른 기울기를 보였다(Figure 3). 이는 동일한 어음명료지수 값에서 K-HINT 문장표를 사용하였을 때 더 높은 어음인지도 값이 산출될 수 있다고 해석할 수 있다. 문장표의 종류에 따른 전달함수 예측 값의 차이는 문장표를 구성하고 있는 예측 가능성의 차이로 설명할 수 있다. K-HINT 문장표(Moon et al., 2008)의 경우에는 100% 높은 예측도(high-predictability) 문장표로만 구성된 반면에 K-SPIN 문장표(An et al., 2002)는 50%의 높은 예측도와 50%의 낮은 예측도(low-predictability) 문장표로 구성되어 K-SPIN 문장표에 비해 K-HINT 문장표가 동일 조건하에서 더 높은 어음인지도로 측정되었을 가능성이 있다. 본 비교는 한국어 문장표의 종류에 따라 주파수중요함수 및 전달함수가 다르게 산출될 수 있음을 보여준다. 따라서 보청기 예후 예측 등의 목적으로 어음명료지수를 활용할 경우에는 검사 자극음의 종류를 확인하고 해당 자극음을 기반으로 산출한 어음명료지수를 사용해야 정확한 예측 값을 얻을 수 있음을 시사한다.

본 연구의 결과는 대상자의 가청능력에 따른 명료어음의 어음인지도 향상 효과에 대한 정보를 제공하고 있다. 이러한 정보는 난청자 및 보청기 사용자의 중재 전략 선택 과정에서 명료어음의 적용에 따른 인지도 이득 효과에 대한 예측 정보를 제공하여 상담 과정에서 명료어음의 적용 여부를 결정하는 데 참고할 수 있다. 예를 들어 보청기 사용자가 보청기 착용 후 조용한 환경에서는 가청능력이 0.45 SII로 약 81% 정도의 어음인지도 수행력을 보였으나, 0 dB SNR 상황에서 가청능력이 0.3 SII로 저하되었고 이는 약 46%의 어음인지도 수행력을 보이는 경우를 가정해 볼 수 있다(Figure 1). 이때 청능사는 보청기 사용자의 대화 상대인 화자가 0 dB SNR과 같이 소음이 있는 상황에서 명료어음을 사용할 경우 명료어음에 의해 산출된 주파수중요함수와 보청기

사용자의 청력 역치에 기반하여 가청능력이 0.35 SII로 향상되고 명료어음 전달함수로 예측된 어음인지도 점수가 약 64%로 명료어음 사용에 따른 어음인지도 향상이 약 18%(= 64 - 46) 정도 될 수 있는 점을 명료어음 어음명료지수를 통하여 예측할 수 있다(Figure 1). 따라서 청능사는 해당 보청기 사용자의 소음 속 어음인지도 향상을 위해서 명료어음이 약 18% 정도 인지도 향상에 기여할 수 있다는 정보를 기반으로 보청기 사용자 및 보호자에게 명료어음을 유용한 의사소통 전략으로 활용할 수 있는 근거를 확보할 수 있으며 예상되는 실제적인 이득을 상담 과정에 포함할 수 있다. 이와 반대로 명료어음 사용에 따른 어음인지도 이득이 적다고 예측된 경우에는 다른 중재 전략을 우선적으로 검토해 볼 수 있는 근거 또한 제공할 수 있을 것이다.

명료어음의 발생법은 많은 연구자들이 다양한 방법으로 제시하고 있다. 명료어음 발생법으로는 정확한 발성을 하기 위해 집중하면서 명확하게 발음하려는 노력이 필요한 방법(Bradlow et al., 2003; Ferguson, 2004), 일반 발생법보다는 입 모양을 크게 만든 후 발음하는 방법(Godoy et al., 2014; Smiljanić & Bradlow, 2007), 주요 문구를 강조해서 발음하되 특히 명사, 동사, 형용사를 강조하는 방법(Caissie et al., 2005; Tye-Murray, 2004) 등의 발생 방법들이 많은 연구자들에 의해 제안되고 있다. 하지만 주파수중요함수의 경우 화자에 대한 음향학적 특징에 따라서 주파수별 특성이 다르게 나타날 수 있다(Healy et al., 2013; Yoho et al., 2018). 심지어 발성의 음향적 발현에 영향을 주는 성대의 공간과 크기에 따라서도 달라진다고 보고된 연구가 있다(Peterson & Barney, 1952). 이는 사용하고자 하는 명료어음 발생법에 따라서 본 연구에서 제시한 주파수중요함수 및 전달함수와는 다른 결과가 도출될 수 있음을 의미한다. 본 연구는 명료어음의 다양한 발생 방법 중 녹음에 참여하는 화자에게 “한국어에 대해 익숙하지 않은 외국인에게 한국어로 이야기하듯이 발생해주세요”라는 안내를 사용하였다(Godoy et al., 2014). 하지만 한국어 명료어음 발생법 중 가장 효과가 높은 발생 방법에 대한 연구는 이루어지지 않았으므로 후속 연구를 통하여 가장 효과적인 명료어음 발생법을 확인하고 어음명료지수를 산출한다면 더욱 효과적인 중재 전략으로 명료어음과 명료어음 어음명료지수를 활용할 수 있을 것이다.

명료어음은 청자의 난청 및 주변 소음 등의 영향으로 청취 환경이 열악할 때 어음인지도를 향상시킬 수 있는 효과적인 의사소통 전략으로 여겨지고 있다. 하지만 명료어음의 어음인지도 이득 효과는 대상자의 가청능력에 따라 다양하게 나타날 수 있다. 본 연구에서는 대상자의 가청능력에 따른 명료어음의 어음인지도 이득 효과를 예측할 수 있는 주파수중요함수와 전달함수를 도출하였다. 본 연구의 결과는 난청자 및 보청기 사용자의 어음인지도 향상을 위한 중재 전략으로 명료어음을 고려

할 때 대상자의 가청능력에 따른 어음인지도 항상 효과 예측에 대한 가이드라인을 제공할 수 있을 것이다.

중심 단어 : 어음명료지수 · 명료어음 · 주파수중요함수 · 전달함수 · 어음인지도.

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.

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.

Author Contributions

I.J. designed and managed this manuscript, and reviewed its final version as the corresponding author. H.O and S.H. worked for data collection and analysis under I.J.'s supervision. Also, H.O. wrote a draft of the manuscript.

ORCID iDs

Hongyeop Oh <https://orcid.org/0000-0002-5493-4648>
In-Ki Jin <https://orcid.org/0000-0002-0834-5981>

REFERENCES

- 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.
- American National Standard Institute. (1997). *ANSI S3.5 (R2012). Methods for Calculation of the Speech Intelligibility Index*. New York, NY: Acoustical Society of America.
- American National Standard Institute. (2013). *ANSI/ASA S3.1-1999 (R2013). Maximum Permissible Ambient Noise Levels For Audiometric Test Rooms*. New York, NY: Acoustical Society of America.
- An, L. J., Kim, J. S., & Pae, S. Y. (2002). The study on developing a test of speech perception in noise. *Journal of Audiology and Otology*, 6(2), 118-127.
- Bradlow, A. R., Kraus, N., & Hayes, E. (2003). Speaking clearly for children with learning disabilities: Sentence perception in noise. *Journal of Speech, Language, and Hearing Research*, 46(1), 80-97.
- 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.
- Caissie, R., Campbell, M. M., Frenette, W. L., Scott, L., Howell, I., & Roy, A. (2005). Clear speech for adults with a hearing loss: Does intervention with communication partners make a difference? *Journal of the American Academy of Audiology*, 16(3), 157-171.
- Ferguson, S. H. (2004). Talker differences in clear and conversational speech: Vowel intelligibility for normal-hearing listeners. *The Journal of the Acoustical Society of America*, 116(4 Pt 1), 2365-2373.
- Fleiss, J. L. (1981). Balanced incomplete block designs for inter-rater reliability studies. *Applied Psychological Measurement*, 5(1), 105-112.
- Godoy, E., Koutsogiannaki, M., & Stylianou, Y. (2014). Approaching speech intelligibility enhancement with inspiration from Lombard and Clear speaking styles. *Computer Speech and Language*, 28(2), 629-647.
- 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.
- Heo, S., Oh, H., & Jin, I. K. (2018). Comparisons of the Korean band-importance and absolute transfer functions according to speakers' gender. *Audiology and Speech Research*, 14(4), 236-242.
- Hinkelmann, K. & Kempthorne, O. (2005). *Design and Analysis of Experiments, Volume 2: Advanced Experimental Design*. Hoboken, NJ: Wiley.
- Jin, I. K., Kates, J. M., & Arehart, K. H. (2017). Sensitivity of the speech intelligibility index to the assumed dynamic range. *Journal of Speech, Language, and Hearing Research*, 60(6), 1674-1680.
- 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.
- Jung, A. F. (1961). Interviewer differences among automobile purchasers. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 10(2), 93-97.
- Kates, J. M. (2013). Improved estimation of frequency importance functions. *The Journal of the Acoustical Society of America*, 134(5), EL459-EL464.
- 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.
- Lee, K., Oh, H., Shin, S., Heo, S., & Jin, I. K. (2018). Comparison of dynamic range of clear speech and conversational speech in Korean. *Audiology and Speech Research*, 14(2), 143-147.
- 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.
- Payton, K. L., Uchanski, R. M., & Braida, L. D. (1994). Intelligibility of conversational and clear speech in noise and reverberation for listeners with normal and impaired hearing. *The Journal of the Acoustical Society of America*, 95(3), 1581-1592.
- Peterson, G. E. & Barney, H. L. (1952). Control methods used in a study of the vowels. *The Journal of the Acoustical Society of America*, 24(2), 175-184.
- Picheny, M. A., Durlach, N. I., & Braida, L. D. (1985). Speaking clearly for the hard of hearing I: Intelligibility differences between clear and conversational speech. *Journal of Speech, Language, and Hearing Research*, 28(1), 96-103.
- Smiljanić, R. & Bradlow, A. R. (2007). Proceedings of the XVI International Congress of Phonetic Sciences: *Clear Speech Intelligibility: Listener and Talker Effects*. Saarbrücken: Saarland University.
- 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.
- Tye-Murray, N. (2004). *Foundations of Aural Rehabilitation: Children, Adults, and Their Family Members*. (2nd ed.), (pp. 41-478). Clifton Park, NY: Thomson Delmar Learning.
- Uchanski, R. M. (2005). Clear speech. In Pisoni, D. B. & Remez, R. E. *The Handbook of Speech Perception* (pp. 207-235). Malden, MA: Wiley-Blackwell.
- Wong, L. L. N., Ho, A. H. S., Chua, E. W. W., & Soli, S. D. (2007). Development of the Cantonese speech intelligibility index. *The Journal of the Acoustical Society of America*, 121(4), 2350-2361.
- Yang, Z. F., Lee, P. C., Chen, W. H., & Leu, J. G. (2011). Extension of structural watermarks based on balanced incomplete block designs. *Journal of Information Hiding and Multimedia Signal Processing*, 2(4), 354-365.
- Yates, F. (1936). Incomplete randomized blocks. *Annals of Human Genetics*, 7(2), 121-140.
- Yoho, S. E., Healy, E. W., Youngdahl, C. L., Barrett, T. S., & Apoux, F. (2018). Speech-material and talker effects in speech band importance. *The Journal of the Acoustical Society of America*, 143(3), 1417-1426.