

# Evaluation of the Korean Matrix Sentence Test: Verification of the List Equivalence and the Effect of Word Position

Kyung Hyun Kim, Jae Hee Lee

Department of Audiology and Speech-Language Pathology, Hallym University of Graduate Studies, Seoul, Korea

한국어 Matrix 문장 검사 도구 평가: 목록 간 동질성과 단어 위치의 영향에 대한 검증

김 경 현 · 이 재 희

한림국제대학원대학교 청각언어치료학과

**Purpose:** Multilingual Matrix Sentence Tests have been developed for reliable, precise, and internationally comparable sentence-in-noise tests. The lists of the Korean Matrix Sentence Test have been constructed and optimized in the previous study. The sentences are generated by a base matrix of 50 Korean words (10 names, 10 adjectives, 10 objects, 10 numerals, and 10 verbs). As suggested by International Collegium of Rehabilitative Audiology working group, this study conducted evaluation tests to verify the list equivalence and the effect of word position at a fixed signal-to-noise ratio (SNR). **Methods:** For experiment I, fifteen normal-hearing (NH) listeners participated. As a material, the Korean Matrix sentences optimized and level-adjusted were used. Twelve test lists of ten sentences were combined to six test lists of twenty sentences (double lists) such that each listener was tested with six double-lists in an open-set format. The percentage of correct responses was measured based on word scoring at fixed SNRs of -10, -8, -6, -4 dB SNRs, with a constant noise level of 65 dB sound pressure level. List-specific intelligibility functions for each list were derived by fitting the percent-correct scores to the logistic regression function, yielding speech reception threshold (SRT) at which 50% intelligibility and also the slope at the SRT across listeners. For experiment II, thirty hearing-impaired (HI) subjects participated and all of them have used hearing aids bilaterally. The aided puretone threshold average of HI listeners was 35.9 dB HL. The materials of the Korean Matrix sentences were presented through loudspeakers while wearing listener's hearing aids. Based on word scoring, the listeners' recognition score was measured in four listening conditions (quiet, 6, 3, 0 dB SNRs). **Results:** In experiment 1, the mean SRT and mean slope across the six test lists were -9.56 dB SNR and 14.03 %/dB, respectively. Across the fixed SNRs, mean difference of SRT was within 1 dB across lists. Based on word scoring, the NH listeners were found to poorly recognize the word at the first position than words at other positions. The mean SRT and slope of the Korean Matrix test were comparable to results of other multilingual Matrix tests. In experiment 2, the results of HI listeners depended on the listening condition (quiet, 6, 3, 0 dB SNR). Regardless of condition, HI listeners poorly recognized words at the first position than words at other positions. The recognition score obtained from the closed-set format was greater than from the open-set test format. **Conclusion:** The equivalence in intelligibility was confirmed by the results of the Korean Matrix sentence recognition at the fixed SNRs. The results obtained from Korean Matrix test were comparable to the results of other multilingual Matrix tests. The optimized and evaluated Korean Matrix Sentence Test can be used for accurate and reliable speech intelligibility in noise measurements.

**Key Words:** Sentence-in-noise recognition, Speech audiometry, Sentence recognition score.

**Received:** April 9, 2018 / **Revised:** April 20, 2018 / **Accepted:** April 23, 2018

**Correspondence:** Jae Hee Lee, Department of Audiology and Speech-Language Pathology, Hallym University of Graduate Studies, 427 Yeoksam-ro, Gangnam-gu, Seoul 06197, Korea

**Tel:** +82-2-2051-4952 / **Fax:** +82-2-3451-6618 / **E-mail:** leejaehee@hallym.ac.kr

## INTRODUCTION

청능사(audiologist)는 난청인의 보장구 선택 및 조절, 청각 재활의 계획 및 시행 등을 위해 순음청력검사 외에 다양한 어

음청각검사를 실시한다. 이때 사용하는 도구는 검사의 목적으로 개발되고 표준화된 도구이어야 한다(Killion et al., 2004). 우리는 일상 생활에서 소음이 있는 환경을 자주 접하므로 대상자의 일상 생활 속 의사소통능력을 정량화하기 위해 소음하 어

음인지도 평가를 시행해야 한다. 국외에서는 Hearing in Noise Test (Nilsson et al., 1994), Quick Speech-in-Noise Test (Killion et al., 2004), Matrix Sentence Test (Akeroyd et al., 2015; Kollmeier et al., 2015) 등 대상자의 반응에 따라 신호대소음비 (signal-to-noise ratio, SNR)를 조절하며 검사하는 도구들을 연구와 임상 진단 시 활용하고 있다. 이외에도 고정된 SNR에서 문장을 제시하고 결과를 백분율로 점수화하는 검사 도구들이 있으나(Speech Perception in Noise Test, Connected Speech Test 등) 검사에서 사용되는 고정된 SNR 조건이 어떤 난청인에게는 너무 쉽거나 혹은 너무 어려울 수 있고, 50% 정도의 인지도를 보이는 SNR을 찾기 위해 여러 개의 목록을 사용하여야 하는 단점이 있다(Lee & Yi, 2017).

이러한 단점을 고려하여 최근에는 청자의 반응에 따라 자극 음 혹은 소음의 강도를 변화하는 변동형(adaptive) 소음하 문장인지검사 도구를 개발하는 추세이다. 특히 기존에 개발된 everyday sentence 문장인지검사 목록은 한정된 문장 수를 포함하므로 대상자가 문장 전체 혹은 문장 내 일부 단어를 외워서 대답할 수 있다(Warzybok et al., 2015). 검사에 대한 훈련 효과(training effect)에 대한 우려를 최소화하면서 반복적인 문장인지검사가 가능한 Matrix 문장인지검사 도구가 다양한 외국어 버전으로 개발되었다(Kollmeier et al., 2015). 현재까지 15개 언어로 Matrix 문장인지검사 도구가 발표되었고(미국 영어, 호주 영어, 영국 영어, 덴마크어, 네덜란드어, 핀란드어, 프랑스어, 독일어, 이탈리아어, 노르웨이어, 폴란드어, 러시아어, 스페인어, 스웨덴어, 터키어), 현재 한국어를 비롯하여 이스라엘어, 이란어, 아랍어, 히브리어, 페르시아어 등의 Matrix 문장인지검사 도구가 개발 중이다(Akeroyd et al., 2015; Lee & Yi, 2017).

Akeroyd et al.(2015)은 다국어 Matrix 문장 목록 개발을 전 세계에서 공통적으로 사용할 수 있도록 가장 첫 단계인 단어 및 문장 구성부터 녹음 화자 선택 및 소음 제작, 문장 최적화, 검사 방법 및 타당도 검사 방법까지 최대한 모든 과정을 동일하게 할 것을 제안하였다. 모든 Matrix 검사 도구는 10 × 5 행렬(matrix)에 포함된 50개의 단어 중 5개의 단어를 무작위로 선택 및 조합하여 문장을 생성한다. 여기서 5개의 단어는 주어, 형용사, 명사, 수사, 동사 각각 10개의 단어 중 하나씩 임의 추출되므로 총 10<sup>5</sup>개의 서로 다른 문장이 생성될 수 있다. 이와 같이 10<sup>5</sup>개의 문장 중 하나가 무작위로 선택되어 제시되므로 대상자가 문장을 외우거나 문장 내 단어를 추측하여 대답하기가 어렵다.

현재 한국어 버전 Matrix 검사 도구 개발을 위해 Akeroyd et al.(2015)이 제안한 방식대로 한국어 Matrix 문장 목록을 구성하고 최적화(optimization)하였다(Yi, 2016). 그러나 소프트웨어에 기반하여 변동형 소음하 문장인지검사를 시행하기 위해

서는 건청인, 난청인을 대상으로 문장 목록의 검증, 소음하 인지도에 영향을 미치는 다양한 요소들에 관한 다양한 후속연구가 진행되어야 한다. 따라서 본 연구에서는 Akeroyd et al.(2015)이 제안한 목록 검증에 대한 두 가지 절차(evaluation 1, 2)를 참고하여 건청인, 난청인을 대상으로 다음의 사항을 평가하고자 하였다. 1) 실험 1에서는 건청인에게 4개의 SNR 조건(-10, -8, -6, -4 dB SNR)에서 6개의 Matrix 문장 목록을 제시한 후 목록 간 동질성(list equivalence) 여부를 검증하고, 각 단어의 인지도(word scoring)를 시행하여 문장 내 단어 위치(word position)에 따른 인지도를 확인하며, 한국어 Matrix 문장인지 결과[speech reception threshold (SRT), slope]와 다른 언어 Matrix 문장인지 결과를 비교하고자 하였다. 2) 실험 2에서는 난청인을 대상으로 3개의 SNR 조건(0, 3, 6 dB SNR)에서 Matrix 문장 목록을 제시하여 평균 인지도뿐 아니라 각 단어의 인지도를 구하여 단어 위치에 따른 인지도를 비교하고자 하였다. 또한 검사 방법(open-set vs.closed-set)에 따른 인지도 차이를 확인하였다.

## MATERIALS AND METHODS

실험 1에서는 건청인 15명을 대상으로, 실험 2에서는 난청인 30명을 대상으로 실험을 진행하였다. 실험 1과 실험 2의 정보를 아래와 같이 구분하여 기술하였다. 모든 대상자는 연구 참여 전 연구의 목적 및 절차에 대한 설명을 듣고 연구 참여에 동의하였다(IRB: HUGSAUD324180).

### Experiment 1

#### 연구 대상

실험 1에 참여한 건청인은 총 15명으로(평균 32.33세, 연령 범위 25~45세, 남 10, 여 5) 이경 검사 결과 외이 및 중이에 이상이 없고 소음에 노출된 경험이 없음을 확인하였다. 건청인의 순음청력검사 결과(GSI 61; Grason-Stadler, Eden Prairie, MN, USA) 모든 대상자의 양이 순음청력역치가 250~8,000 Hz 범위에서 20 dB HL 미만이었다.

#### 연구 절차

검사에 사용한 Matrix 문장은 선행연구(Yi, 2016)에서 각 단어 음원의 레벨을 수정(level adjustment)하여 최적화를 완료한 음원이었다. 한국어 Matrix 문장 목록의 평가를 위하여 Akeroyd et al.(2015)이 제안한 대로 10개의 Matrix 문장으로 구성된 12개의 문장 목록을 20개의 문장으로 구성된 6개의 문장 목록으로 합쳐서 사용하였다.

모든 실험은 국제표준에서 권고하는 소음허용 수준(Interna-

tional Organization for Standardization, 2012)을 만족하는 방음실에서 실시하였고, 헤드폰(HDA 200; Sennheiser Electronic GmbH & Co. KG, Wedemark, Germany), Fireface UCX 디지털-아날로그 변환기(RME, Haimhausen, Germany), 컴퓨터(Sens-9, Samsung electronics, Suwon, Korea)를 사용하여 문장 음원을 제시하였다.

앞서 설명한 것처럼 실험 1의 연구 목적은 목록 간 동질성 여부를 확인하고 문장 내 단어 위치에 따른 인지도를 확인하며, 다른 언어 Matrix 검사 결과와 비교하는 것이었다. 따라서 선행 연구(Hochmuth et al., 2012; Puglisi et al., 2015)에서 20~80% 인지도를 보일 것으로 예측하고 사용한 SNR 범위를 듣기 조건(-9.5~-4 dB)으로 사용하였다. 한국어 Matrix 문장의 어음 스펙트럼을 가지도록 제작한 비유동성(stationary) 소음(Yi, 2016)을 고정된 레벨인 65 dB sound pressure level에서 제시하고 -10, -8, -6, -4 dB SNR 조건에서 목록별 문장인지를 측정하였다.

Matrix 검사 도구의 개발자들은 훈련 효과를 최소화하기 위해 최소한 2개의 문장 목록을 통해 친숙화 과정을 거친 후 실제 검사를 진행할 것을 권장한다(Akeroyd et al., 2015). 따라서 훈련 효과의 영향을 최소화하기 위해 2개의 문장 목록을 듣고 맞추는 연습 과정을 거친 후 실제 실험을 시작하였다. 실험 시 문장 목록과 SNR 제시 순서는 무작위로 선정하여 대상자 간 서로 다르게 하였으며, 측정 시간은 대략 60분이었다. 목록 간 평균 문장인지와 더불어 문장 내 단어 위치에 따른 인지도 차이를 비교하고자 문장 내 각 단어별 인지도와 문장별 인지도를 모두 점수화하였다.

## Experiment 2

### 연구 대상

실험 2에는 난청인 30명이 참여하였다(평균 61세, 연령 범위 24~83세, 남 15, 여 15). 대상자 모두 난청 외에 어지러움, 이명, 신경학적 병력을 가지고 있지 않았다. 순음청력검사 결과(GSI 61; Grason-Stadler) 난청인 대상자의 250~8,000 Hz 이내 옥타브 단위 주파수별 순음청력역치는 오른쪽 귀의 경우 50, 51, 53, 58, 65, 74 dB HL이었고(표준편차: 18.7, 16.3, 12.0, 11.5, 15.4, 17.1), 왼쪽 귀의 경우 48, 51, 53, 58, 66, 74 dB HL(표준편차: 19.4, 15.7, 12.1, 9.4, 11.6, 17.2)이었다. 대상자 모두 양이에 보청기를 사용 중이었다. 양이 보청기를 착용하고 250~4,000 Hz 이내 옥타브 단위의 주파수별 순음청력역치는 34, 35, 36, 37, 41 dB HL이었고, 양이 보청기 착용 후 측정된 500, 1,000, 2,000 Hz 주파수의 평균순음역치(puretone threshold average)는 35.9 dB HL이었다(범위: 25~48 dB HL, 표준편차: 5.5).

60세 이상 대상자의 경우 한국판 간이정신상태검사(Korean version of Mini-Mental State Examination) (Park & Kwon, 1989)를 시행하여 인지 기능의 저하가 없음을 확인하였다.

### 연구 절차

실험 2에서는 난청인을 대상으로 하므로 실험 1에서 건청인에게 제시한 SNR 범위인 -10, -8, -6, -4 dB SNR 조건보다 더 쉬운 듣기 조건을 사용하여야 했다. 난청인을 대상으로 한 선행연구(Kim et al., 2017) 결과를 참고하여 0, 3, 6 dB SNR과 소음이 없는 조건에서 총 4가지 듣기 조건에서 문장 목록을 80 dB SPL에서 제시하여 난청인 대상자의 문장인지를 측정하기로 하였다. 실험 2에서는 난청인에게 고정된 SNR에서 Matrix 문장 검사를 시행하고 전반적인 난이도를 확인하여 단어의 위치에 따른 인지도 차이를 확인하고자 하였다. 실험 1과 마찬가지로 실험 2에서도 실험을 시작하기 전 친숙화 과정을 거친 후 실제 실험을 시작하였다. 실험 시 가장 쉬운 조건인 소음 없는 조건을 가장 먼저, 그 다음 쉬운 순서인 6, 3, 0 dB SNR 조건의 순서로 문장을 제시하였다.

30명의 난청인 대상자 중 15명의 대상자를 무작위로 선택하여 문장 제시 방법(open-set vs. closed-set)에 따른 차이를 보기 위한 평가를 추가로 시행하였다. Matrix 문장인지검사는 10 × 5의 행렬에 포함되는 50개의 단어를 보기로 보여주고 closed-set 검사를 시행할 수도 있고 보기 없이 들은 문장을 따라하게 하는 open-set 검사로도 시행할 수 있다. 검사 도구의 최적화 과정 및 타당도, 신뢰도를 구하는 대부분의 실험에서는 open-set로 들은 문장을 따라 말하는 검사 방법을 따른다. 그러나 연구자들은 open-set 검사가 불가능한 대상자의 경우 closed-set로 검사를 시행해야 할 수 있으므로 두 가지 방법에 따른 차이를 확인할 필요가 있다고 밝혔다(Kollmeier et al., 2015). 따라서 난청인 15명에게 open-set 검사가 끝난 후 closed-set 검사를 추가로 시행하여 그 차이를 확인하였다. 보기를 제시하는 closed-set 검사를 먼저 시행할 수 없으므로 항상 open-set 검사를 시행하였으나 이어서 검사를 진행하지 않고 한 시간가량 휴식시간을 가진 후 closed-set 검사를 시행하였다. 실험 2에 소요된 시간은 대략 60~90분이었다.

### 통계 분석

수집된 결과는 SPSS version 20.0 (IBM Corp., Armonk, NY, USA) 소프트웨어를 이용하여 분석하였고, 모든 통계분석은 유의수준 0.05를 기준으로 하였다. 반복 측정 분산분석을 시행하였을 때 구형성 가정이 만족되지 않았을 경우 Greenhouse-Geisser 수정된 검정 통계량을 보고하였고, Pearson 상관분석을 통해 선형 관계 정도를 분석하였다.

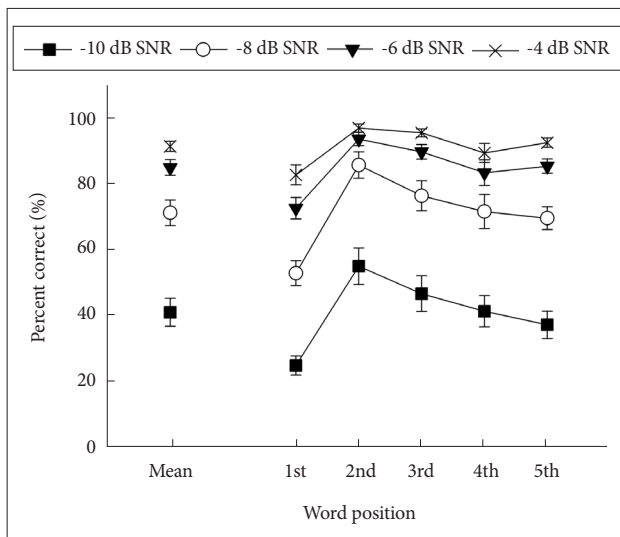
## RESULTS

### Experiment 1

실험 1에서는 건청인 15명에게 -4, -6, -8, -10 dB SNR의 고정된 듣기 조건에서 Matrix 문장 목록 1~6을 제시하여 소음하 문장인지를 측정하였다. -4, -6, -8, -10 dB SNR 조건별 건청인의 평균 문장인지는  $91.24 \pm 5.88\%$ ,  $84.78 \pm 9.32\%$ ,  $71.09 \pm 15.03\%$ ,  $40.91 \pm 16.32\%$ 였다.

Figure 1에 SNR별로 측정된 건청인의 평균 인지도와 단어의 위치에 따른 인지도를 함께 제시하였다. 가장 듣기 쉬운 조건이었던 -4 dB SNR 조건의 경우, 문장 내 각 단어의 순서대로(첫 번째부터 다섯 번째까지) 나열하면 평균 인지도가 82.56%, 96.78%, 95.33%, 89.22%, 92.33%였다(표준편차: 11.56, 4.86, 4.72, 11.09, 5.41). -6 dB SNR 조건에서 위와 동일한 단어의 위치대로 나열하면 평균 인지도가 72.44%, 93.44%, 89.56%, 83.22%, 85.22%였고(표준편차: 12.60, 7.62, 8.72, 15.09, 8.28), -8 dB SNR에서 측정된 경우 동일 순서대로 나열하면 52.78%, 85.56%, 76.22%, 71.44%, 69.44%였다(표준편차: 14.44, 15.36, 17.66, 20.07, 13.27). 마지막으로 가장 듣기 어려운 조건이었던 -10 dB SNR 조건의 경우 문장 내 단어 위치에 따른 단어인지는 24.78%, 54.89%, 46.56%, 41.22%, 37.11%였다(표준편차: 11.14, 21.43, 20.95, 18.34, 16.12).

문장 내 첫 번째 단어의 인지도가 다른 위치의 단어인지도보다 낮은 경향이 있었으므로 이를 확인하기 위해 SNR과 단어 위치(word position)를 그룹 내 독립변수로 사용하여 반복 측정된 이원변량분산분석을 시행하였다. 분석 결과, SNR 조건에 따른 주 효과가 유의하였다( $F(1.22, 17.12) = 204.71, p < 0.05$ ).

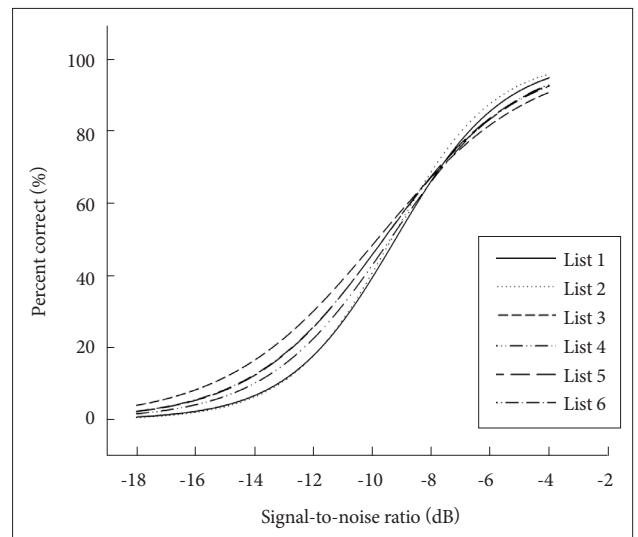


**Figure 1.** Mean percent-correct score of normal-hearing listeners when measured from -10 dB SNR to -4 dB SNR (error bar: standard error). SNR: signal-to-noise ratio.

Bonferroni 수정에 의한 다중비교 결과 네 개의 SNR 조건 간 문장인지도가 모두 유의하게 달라 -4 dB SNR보다는 -6 dB SNR의 인지도가 낮았고, -6 dB SNR보다는 -8 dB SNR의 인지도가, -8 dB SNR보다는 -10 dB SNR의 인지도가 유의하게 낮았음을 확인하였다. 단어 위치에 따른 주 효과( $F(4, 56) = 54.70, p < 0.05$ ) 또한 유의하였으며, 다중비교에 의한 대응별 비교 결과 문장 내 첫 단어의 인지도가 다른 위치의 단어인지도에 비해 유의하게 낮았고, 두 번째 단어의 인지도가 다른 위치의 단어인지도에 비해 유의하게 높았고, 세 번째 단어의 인지도가 다섯 번째 단어의 인지도보다 높았다. Figure 1을 통해서 확인할 수 있듯이 듣기 어려운 조건일수록 문장 내 첫 번째로 제시된 단어를 인지하는 능력이 보다 더 저하되었는데, 이러한 경향은 SNR과 단어 위치의 두 변수 간 유의한 이원 상호작용을 통해서도 확인하였다( $F(4.63, 63.84) = 4.30, p < 0.05$ ).

목록 간 동질성을 확인하기 위해 로지스틱 회귀분석(logistic regression)을 이용하여 각 SNR에서 측정된 인지도로부터 심리음향기능곡선을 구하였다. 각 SNR 조건에서 측정된 목록별 심리음향기능곡선은 Figure 2와 같다. 목록별 심리음향기능곡선으로부터 50% 인지도를 보이는 SRT와 SRT를 점점으로 하는 기울기를 구하였고 목록별 평균 SRT와 기울기는 Table 1에 기입하였다.

표를 통해 알 수 있듯이, 목록 1을 제시하였을 때 50%의 인지도를 보이는 SRT는  $-9.20 \pm 1.14$  dB SNR이었고, 목록 2의 SRT는  $-9.28 \pm 1.01$  dB SNR, 목록 3은  $-9.93 \pm 1.74$  dB SNR, 목록 4는  $-9.38 \pm 1.38$  dB SNR, 목록 5는  $-9.75 \pm 1.60$  dB SNR, 목록 6은  $-9.84 \pm 1.99$  dB SNR이었다. 목록 간 총 평균 SRT 값은  $-9.56 \pm 1.50$  dB SNR이었다. 선행연구(Yi, 2016)에



**Figure 2.** List-specific intelligibility functions for normal-hearing listeners.

**Table 1.** Mean SRT and slope across lists obtained from Korean Matrix sentence-in-noise measurements at fixed SNRs

	List 1	List 2	List 3	List 4	List 5	List 6	Mean
SRT (dB SNR)	-9.20 (1.14)	-9.28 (1.01)	-9.93 (1.74)	-9.38 (1.38)	-9.75 (1.60)	-9.84 (1.99)	-9.56 (1.50)
Slope (%/dB)	16.33 (5.38)	17.36 (4.89)	10.85 (2.86)	13.85 (3.76)	12.86 (4.10)	12.94 (4.33)	14.03 (4.72)

SRT: speech reception threshold, SNR: signal-to-noise ratio

**Table 2.** Mean SRT and slope across lists obtained from multilingual Matrix sentence-in-noise measurements at fixed SNRs

	Korean	Polish Matrix (2010)	Finnish Matrix (2014)	Italian Matrix (2014)	Dutch Matrix (2014)	Turkish Matrix (2015)	Russian Matrix (2015)
Mean SRT (dB SNR)	-9.56	-9.6	-10.1	-7.3	-8.4	-8.3	-9.5
Mean slope (%/dB)	14.03	17.1	16.7	13.3	10.2	14.1	13.8

SRT: speech reception threshold, SNR: signal-to-noise ratio

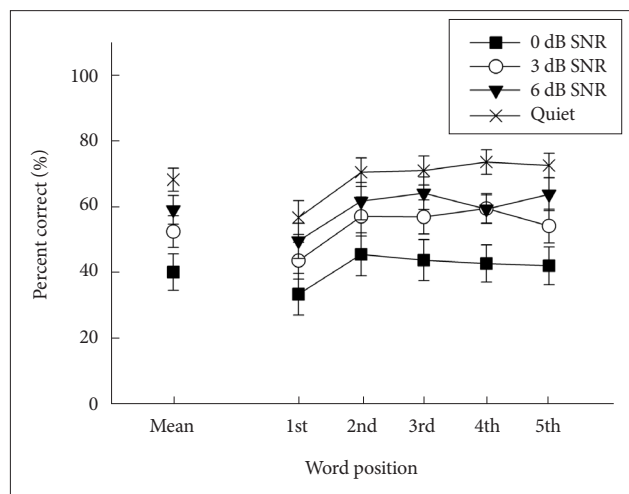
서 482개의 어휘 음원을 최적화한 후 측정한 SRT가  $-9.81 \pm 0.26$  dB SNR이었음을 고려하면 약 0.3 dB 이내의 차이를 보이는 수치였다. Table 1에 제시하였듯이 SRT를 점점으로 하는 기울기를 목록 평균화한 값은  $14.03 \pm 4.72$  %/dB이었으며, 목록 별로 기울기는 목록 1의 경우  $16.33 \pm 5.38$  %/dB, 목록 2는  $17.36 \pm 4.89$  %/dB, 목록 3은  $10.85 \pm 2.86$  %/dB, 목록 4는  $13.85 \pm 3.76$  %/dB, 목록 5는  $12.86 \pm 4.10$  %/dB, 목록 6은  $12.94 \pm 4.33$  %/dB이었다.

위에서 측정한 한국어 Matrix 문장인지 결과(SRT, slope)와 다른 언어 Matrix 문장인지 결과를 비교하고자 2010년 이후에 출판된 Matrix 문장인지 연구 결과와 비교하였다. Table 2는 한국어 Matrix 인지 결과와 2010년 이후에 발표된 폴란드어 (Ozimek et al., 2010), 핀란드어(Dietz et al., 2014), 이탈리아어 (Puglisi et al., 2015), 네덜란드어(Houben et al., 2014), 터키어 (Zokoll et al., 2015), 러시아어(Warzybok et al., 2015) Matrix 문장인지 결과를 보여준다. 표를 통해 알 수 있듯이, 한국어 Matrix 문장을 이용하여 측정한 SRT 평균값은  $-9.56$  dB SNR, SRT 범위는  $-9.93 \sim -9.20$  dB SNR이었다. 목록 간 SRT 값이  $\pm 1$  dB 이내로 유사하였음을 알 수 있다. 한국어 Matrix 문장인지 결과에서 SRT에 접점하는 기울기 평균값은  $14.03$  %/dB이었고, 기울기 범위는  $10.85 \sim 17.36$  %/dB이었다.

**Experiment 2**

실험 2에서는 난청인을 대상으로 세 개의 소음하 조건(0, 3, 6 dB SNR)과 소음이 없는 조건(quiet)에서 Matrix 문장 목록을 제시하여 평균 인지도를 측정하였다. 소음하 조건별로 문장 내 단어 위치에 따른 인지도에 차이가 있는지 확인하고자 단어 위치에 따른 인지도 또한 함께 비교하였다. 또한 문장 제시 방법에 따라(open-set vs. closed-set) 난청인의 인지도가 달라지는 지 등을 평가하였다.

먼저, 듣기 조건별로 소음하 문장인지도를 측정한 후 난청인의 평균 문장인지도뿐 아니라 문장 내 단어의 위치에 따른 인



**Figure 3.** Mean percent-correct score of hearing-impaired listeners obtained from quiet and 0, 3, 6 dB SNR conditions (error bar: standard error). SNR: signal-to-noise ratio.

지도를 함께 Figure 3에 나타내었다. 세 개의 소음하 조건(0, 3, 6 dB SNR)과 소음이 없는 조건(quiet)에서 측정한 평균 문장 인지도는  $39.97 \pm 30.38\%$ ,  $52.27 \pm 26.36\%$ ,  $58.92 \pm 23.79\%$ ,  $68.10 \pm 19.04\%$ 였다. 각 조건별로 단어 위치에 따른 인지도를 분석한 결과, 제시 조건 중 가장 어려운 조건인 0 dB SNR에서 측정한 인지도는 단어 위치 순서대로  $33.28\%$ ,  $45.35\%$ ,  $43.62\%$ ,  $42.59\%$ ,  $41.90\%$ (표준편차: 34.29, 35.18, 33.54, 30.40, 31.01)로 개인 간 편차가 매우 컸음을 알 수 있다. 3 dB SNR에서 측정한 평균 인지도는 위와 동일한 단어 위치 순서대로  $43.45\%$ ,  $56.90\%$ ,  $56.72\%$ ,  $59.31\%$ ,  $53.97\%$ 였고(표준편차: 30.00, 32.08, 27.91, 24.48, 25.76), 6 dB SNR 조건에서 측정한 평균 인지도를 위의 단어 위치 순서대로 나열하면  $49.48\%$ ,  $61.55\%$ ,  $63.97\%$ ,  $59.14\%$ ,  $63.62\%$ 였다(표준편차: 28.77, 30.68, 26.57, 22.96, 26.82). 마지막으로 소음이 없는 조건에서 측정한 단어 위치별 평균 인지도는 위치 순서대로  $56.55\%$ ,  $70.35\%$ ,  $70.86\%$ ,  $73.45\%$ ,  $72.41\%$ 였다(표준편차: 27.71, 23.45, 23.57, 20.09, 20.12).

실험 1에서 확인한 결과와 마찬가지로 실험 2의 난청인 결과

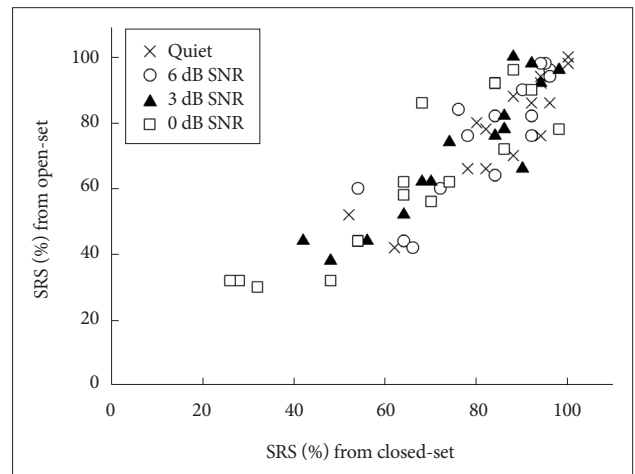
를 통해 듣기 조건(0, 3, 6 dB SNR, quiet)과 단어 위치를 그룹 내 독립변수로 사용하여 반복 측정된 이원변량분산분석을 시행하였다. 먼저, 듣기 조건이 인지도에 미치는 주 효과는 유의하였고[F(1.40, 2.35) = 28.88,  $p < 0.05$ ], Bonferroni 수정에 의한 다중비교 결과 6 dB SNR 조건과 소음이 없는 조건 간 인지도는 유의하게 다르지 않았으나, 나머지 조건 간 결과는 유의하게 달랐다. 단어 위치에 따른 주 효과 역시 유의하였으며[F(2.35, 65.65) = 7.19,  $p < 0.05$ ] 대응비교 결과 실험 1의 건청인 결과와 유사하게 난청인의 결과에서도 문장 내 첫 단어의 인지도가 나머지 단어들을 인지한 결과에 비해 유의하게 낮았다. 그 외 두 번째~다섯 번째 단어들 간 결과는 차이가 없었다. SNR과 단어 위치의 두 변수 간 이원 상호작용은 유의하지 않았으므로 [F(4.37, 122.38) = 0.98,  $p > 0.05$ ] 듣기 조건에 상관없이 난청인은 문장 내 다른 위치의 단어들보다 첫 단어를 인지하는 데 어려움을 가졌음을 알 수 있다.

추가로 문장 제시 방법(open-set vs. closed-set), 듣기 조건(0, 3, 6 dB SNR, quiet)을 두 개의 그룹 내 독립변수로 사용하여 문장인지도(종속변수)에 유의한 영향을 주는지 반복 측정된 이원변량분산분석을 시행하였다. 분석 결과 보기를 제시하지 않고(open-set로 검사 시) 소음이 없는 조건과 6, 3, 0 dB SNR 조건에서 평균 인지도는  $78.13 \pm 16.70\%$ ,  $76.4 \pm 18.63\%$ ,  $70.93 \pm 20.59\%$ ,  $61.47 \pm 23.71\%$ 였다. 동일한 듣기 조건(quiet, 6, 3, 0 dB SNR)에서 보기를 미리 보여주고(closed-set) 측정된 평균 인지도는  $85.47 \pm 13.6\%$ ,  $82.2 \pm 13.34\%$ ,  $76 \pm 17.44\%$ ,  $65.07 \pm 23.38\%$ 였다.

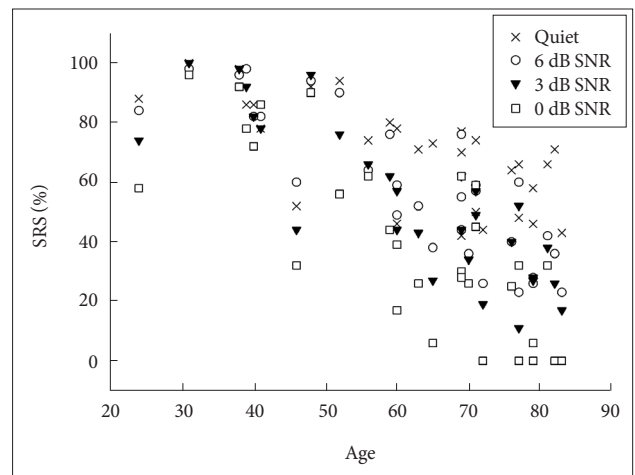
분석 결과, 보기 유무(open-set vs. closed-set)에 따른 주 효과[F(1, 14) = 8.47,  $p < 0.05$ ]와 듣기 조건에 따른 주 효과 모두 유의하였다[F(1.56, 21.85) = 27.22,  $p < 0.05$ ]. 대응 비교 결과 듣기 조건 중 6 dB SNR과 소음이 없는 검사 결과 간 차이를 제외하고는 모두 유의미한 차이를 보였다. 보기 유무와 듣기 조건의 두 변수 간 이원 상호작용은 유의하지 않았으므로[F(1.59, 22.23) = 0.72,  $p > 0.05$ ] 듣기 조건에 상관없이 보기가 없는 open-set 보다는 보기가 제시된 closed-set 검사 시 더 높은 인지도를 보였음을 알 수 있다.

Pearson 상관분석을 통해 듣기 조건별로 open-set와 closed-set 방법에 따른 결과 간 상관성을 확인하였다. Figure 4에서 알 수 있듯이, 모든 조건에서 open-set와 closed-set 결과 간 유의한 상관성이 있었고 높은 양의 상관계수를 보였다(quiet:  $r = 0.88$ , 6 dB SNR:  $r = 0.84$ , 3 dB SNR:  $r = 0.91$ , 0 dB SNR:  $r = 0.89$ ,  $p < 0.05$ ). 따라서 open-set 검사에서 높은 인지도를 보인 사람은 듣기 조건에 상관없이 closed-set 검사에서도 높은 인지도를 보였다.

마지막으로 난청인 전체 대상자( $n = 30$ )의 연령 혹은 청력역



**Figure 4.** Scatterplot of SRS between closed-set and open-set test formats. SRS: sentence recognition score, SNR: signal-to-noise ratio.



**Figure 5.** Scatterplot of relations between listeners' age and SRS. SRS: sentence recognition score, SNR: signal-to-noise ratio.

치에 따라 각 듣기 조건에서 측정된 문장인지도와 관련성을 보이는지 Pearson 상관분석을 시행하였다. Figure 5에 제시하였듯이, 모든 듣기 조건에서 보기 없이 측정된 문장인지도 결과가 대상자의 나이와는 유의한 상관성을 보였고(quiet:  $r = -0.67$ , 6 dB SNR:  $r = -0.84$ , 3 dB SNR:  $r = -0.83$ , 0 dB SNR:  $r = -0.78$ ,  $p < 0.05$ ), 보청기를 착용하지 않았을 때의 평균 순음역치 혹은 보청기 착용 시의 평균 순음역치와는 유의한 관련성이 없었다( $p > 0.05$ ).

## DISCUSSIONS

국내의 경우 표준화된 소음하 어음청각검사 도구가 부족하여 난청인의 소음하 의사소통능력을 측정하는 데 한계가 있다. 따라서 국내 임상 현장에서는 소음 제시 없이 단어인지도(word recognition score)를 주로 측정하고 있는데, 이 결과를 통해 난

청인의 일상 생활 속 의사소통능력을 파악하거나 소음하 상황에서 보장구 혜택 정도를 측정하기는 어렵다.

최근에 개발된 Matrix 소음하 문장인지검사는 소프트웨어에 기반하여 SNR을 조절하는 변동형 검사 도구이다. Akeroyd et al.(2015)은 다국어 Matrix 문장 검사 도구 개발을 위해 문장 목록 구성, 화자 선택, 소음 제작, 문장 최적화 및 검사 방법, 타당도 검사 등에 대한 자세한 가이드라인을 발표하여, 서로 다른 연구자들이 개발하더라도 다국어 검사 결과를 비교할 수 있도록 하였다. 현재 15개의 언어로 Matrix 검사 도구가 개발되었고, 한국어 버전의 경우 한국어 Matrix 문장 목록 구성 및 최적화 과정까지 진행되었다(Yi, 2016). 본 연구에서는 Akeroyd et al.(2015)이 제안한 검증 절차 중 하나로 고정된 SNR에서 목록 간 동질성을 확인하고 그 외에 단어 위치, 검사 방법 등에 따른 인지도를 확인하여 다국어 Matrix 검사 결과와 비교하고자 하였다.

실험 1에서는 목록 간 동질성을 확인하기 위해 12개의 문장 목록(목록당 10개 문장)을 6개의 목록으로 결합하여(목록당 20개 문장) -4, -6, -8, -10 dB SNR에서 모든 목록을 제시하였다. 결과로는 건청인의 목록별 심리음향기능곡선을 구하고 50% 인지도에 해당하는 SRT, SRT에 접점하는 기울기를 구하였다. 실험 1의 측정 결과, 건청인이 50%의 인지도를 보인 평균 SRT는 -9.56 dB SNR, 평균 기울기는 14.03 %/dB이었으며, 목록 간 심리음향기능곡선 또한 유사하였다. 14개 언어의 Matrix 문장인지 검사 결과를 비교한 Kollmeier et al.(2015)은 서로 다른 언어임에도 불구하고 동일한 문장 구조와 개발 방식을 따랐기 때문에 SRT 범위가 -10.1~-6.0 dB SNR 이내라고 보고하였다. 한국어 Matrix 문장을 이용하여 측정한 평균 SRT가 -9.56 dB SNR(범위: -9.93~-9.20 dB SNR)임을 고려하면 위의 다국어 Matrix SRT 결과 범위 내에 한국어 SRT가 포함됨을 알 수 있다.

Kollmeier et al.(2015)은 14개 언어의 Matrix 인지도를 측정한 결과 평균 기울기가 10.2~17.1 %/dB 이내에 포함되었다고 밝혔다. 한국어 Matrix 문장을 이용하여 측정한 평균 기울기 14.03 %/dB, 범위는 10.85~17.36 %/dB(평균은 14.03 %/dB)이었으므로 위의 다국어 Matrix 기울기 범위 안에 한국어 기울기가 포함됨을 알 수 있다. 다국어 Matrix 결과 중 가장 낮은 SNR 결과(약 -10 dB SNR)를 보인 경우가 Russian Matrix (Warzybok et al., 2015)와 Polish Matrix (Ozimek et al., 2010)였고, 가장 높은 SNR 결과(약 -6.0 dB SNR)를 보인 경우가 Spanish Matrix (Hochmuth et al., 2012)와 French Matrix (Jansen et al., 2012) 검사 도구였다. 이와 같이 Spanish Matrix 혹은 French Matrix 문장의 경우 다른 언어에 비해 SRT와 차이가 났던 이유 중 하나로 언어 특성상 고주파수 음소 개수에 따라 영향을 받았을 수 있다고 하였다(Kollmeier et al., 2015).

실험 2에서는 난청인을 대상으로 0, 3, 6 dB SNR의 소음하

조건에서 Matrix 문장인지도를 측정하였다. 측정 결과 0, 3, 6 dB SNR 조건에서 각각 평균 39.97%, 52.27%, 58.92%의 인지도를 보였다. Polat et al.(2016)은 30명의 인공와우 착용자를 대상으로 Polish Matrix 문장 목록을 사용하여 소음하 문장인지도를 구한 결과, -5, 0, 5 dB SNR 소음하 조건에서 측정한 결과 평균 30.23%, 74.37%, 89.30%였다. 본 연구와 마찬가지로 개인 간 편차가 컸으며 인공와우 착용 후 평균 순음역치와 소음하 문장인지도 간 상관성이 없다고 밝혔다. Polat et al.(2016)은 고정된 SNR뿐 아니라 변동형 검사 방법을 통해서도 Matrix 문장 검사를 시행하였다. 소음을 정면 스피커를 통해서 제시한 경우 50% 인지도를 보이는 SRT가 -0.62 dB SNR, 소음을 180°에 위치한 스피커로 제시한 경우 SRT가 -4.9 dB SNR이었다고 하였다. Finnish Matrix 문장 목록을 사용한 Dietz et al.(2014)은 인공와우 착용자들의 평균 SRT가 -3.5 dB SNR이었다고 하였고, Hey et al.(2014)은 독일 인공와우 착용자들의 평균 SRT가 -2 dB SNR이었다고 하였다. 위의 국외 연구 결과는 모두 고정된 SNR이 아닌 변동형 검사를 통해 구한 결과이므로 향후 연구에서도 국내 난청인을 대상으로 변동형 검사 방법을 통한 신뢰도, 타당도 검증 연구가 지속되어야겠다.

모든 Matrix 문장은 동일한 문법 구조를 가지며 동일한 위치에서 해당 품사 단어가 포함되어 있다. Matrix 문장의 경우 문장 내 문맥적 힌트로 단어 5개를 모두 유추하여 답하기는 어렵지만, 대상자가 Matrix 문장의 구조를 파악하고 문장 내 5개의 단어가 있으므로 각 단어를 모두 듣기 위해 소음이 커질수록 보다 집중하여 인지할 가능성이 있다(Akeroyd et al., 2015; Wardenga et al., 2015). 본 연구에서는 Akeroyd et al.(2015)의 가이드라인을 참고하여 각 단어별 인지도를 계산하여 단어 위치에 따라 인지도를 비교하였다. 분석 결과, 건청인과 난청인 공통적으로 소음하 Matrix 문장 중 첫 번째에 위치하는 단어를 인지하는 것을 다른 위치의 단어들을 인지하는 것보다 더 어려워했다. 특히 건청인의 경우 소음의 강도가 더 커질수록 이러한 경향이 더 두드러졌다. 첫 번째 위치한 단어가 “주어”였는데 실제로 문장 내 단어를 최적화하는 과정에서는 각 주어에 대한 인지도는 유의하게 다르지 않았다(Yi, 2016). Neidleman et al.(2015)은 배경소음이 있든 없든 5개의 단어 위치에 따라 단어 회상능력을 비교한 결과 청년, 중장년 대상군 모두 첫 번째 단어들을 회상하는 것을 가장 어려워했다고 보고하였다. 특히 소음이 제시되었을 때 이러한 경향은 청년보다 중장년층에서 더욱 심하였다고 하였다. 따라서 Matrix Sentence Test가 단어 회상능력을 직접적으로 측정하는 도구는 아니지만 검사 시 소음 레벨의 증가에 따라 연령별 단어인지도를 측정하고 이를 단어 위치에 따라 분석하는 연구를 지속할 필요가 있다.

선행연구 중 Hochmuth et al.(2012)과 Puglisi et al.(2015)은

변동형 방법(adaptive method)을 통해 보기가 있을 때(closed-set)와 없을 때(open-set) Matrix 문장인지검사 결과를 비교하였다. 확인 결과 보기가 있을 때 약 0.7~1 dB SNR의 혜택이 있다고 하였다. 이러한 검사 결과는 대부분 건청인을 대상으로 한 결과이므로 본 연구에서는 실험 2에 참여한 난청인 중 15명을 대상으로 한 시간 간격을 두고 조사하였다. 확인 결과 소음이 없는 듣기 조건에서 보기가 주어졌을 때 보기가 없는 open-set 검사 시보다 평균 인지도가 약 7%가량 높았고, 6 dB SNR과 3 dB SNR에서는 약 5%가량 높았고, 0 dB SNR에서는 4%가량 인지도가 더 높았다. 상관분석에서 모든 SNR에서 보기가 있는 검사와 보기가 없는 검사 결과 간 유의한 관련성이 있었으므로 보기를 주지 않는 open-set Matrix 문장인지검사가 불가능한 대상자의 경우 보기를 제시하는 closed-set 방법으로 변경하여 검사를 시행할 수 있겠다. 추후 Matrix 검사 결과를 보고할 때는 어떠한 검사 방법을 통해 시행했는지 반드시 기록하여 검사 기관 간 결과를 비교하는 데 용이하도록 해야겠다.

향후 연구에서는 소프트웨어를 이용한 변동형 검사를 훈련 효과, 화자, 소음 종류 등이 미치는 영향에 대해 확인해야 하며, 검사자, 기관, 지역 간 결과를 비교하여야 하겠다. 이러한 후속 연구로 Matrix 검사 도구의 표준화를 완료한다면 문법적으로 구조가 동일하고 문장의 의미 예측이 불가능한 한국어 Matrix 문장인지검사 도구를 임상 및 연구 목적으로 유용하게 사용할 수 있을 것으로 생각한다.

**중심 단어** : 소음하 문장인지 · 어음청각검사 · 문장인지도.

### Acknowledgments

본 연구의 데이터 수집 및 분석을 위해 도움을 주신 김나경, 이동운 연구원에게 감사드립니다. 이 논문은 2016년 대한민국 교육부와 한국연구재단의 지원을 받아 수행된 연구입니다(NRF-2016S1A5A8020353).

### REFERENCES

- Akeroyd, M. A., Arlinger, S., Bentler, R. A., Boothroyd, A., Dillier, N., Dreschler, W. A., et al. (2015). International Collegium of Rehabilitative Audiology (ICRA) recommendations for the construction of multilingual speech tests. ICRA working group on multilingual speech tests. *International Journal of Audiology*, 54 Suppl 2, 17-22.
- Dietz, A., Buschermöhle, M., Aarnisalo, A. A., Vanhanen, A., Hyyrynen, T., Aaltonen, O., et al. (2014). The development and evaluation of the Finnish Matrix Sentence Test for speech intelligibility assessment. *Acta Oto-Laryngologica*, 134(7), 728-737.
- Hey, M., Hocke, T., Hedderich, J., & Müller-Deile, J. (2014). Investigation of a matrix sentence test in noise: Reproducibility and discrimination function in cochlear implant patients. *International Journal of Audiology*, 53(12), 895-902.
- Hochmuth, S., Brand, T., Zokoll, M. A., Castro, F. Z., Wardenga, N., & Kollmeier, B. (2012). A Spanish matrix sentence test for assessing speech reception thresholds in noise. *International Journal of Audiology*, 51(7), 536-544.
- Houben, R., Koopman, J., Luts, H., Wagener, K. C., van Wieringen, A., Verschuur, H., et al. (2014). Development of a Dutch matrix sentence test to assess speech intelligibility in noise. *International Journal of Audiology*, 53(10), 760-763.
- International Organization for Standardization. (2012). *ISO 8253-3:2012. Acoustics - Audiometric Test Methods - Part 3: Speech Audiometry*. Geneva: International Organization for Standardization.
- Jansen, S., Luts, H., Wagener, K. C., Kollmeier, B., Del Rio, M., Dauman, R., et al. (2012). Comparison of three types of French speech-in-noise tests: A multi-center study. *International Journal of Audiology*, 51(3), 164-173.
- Killion, M. C., Niquette, P. A., Gudmundsen, G., Revit, L. J., & Banerjee, S. (2004). Development of a quick speech-in-noise test for measuring signal-to-noise ratio loss in normal-hearing and hearing-impaired listeners. *The Journal of the Acoustical Society of America*, 116(4 Pt 1), 2395-2405.
- Kim, N. K., Yi, D., Bahng, J., & Lee, J. H. (2017). Study on the reliability and the validity for the sentence-in-noise scores at the fixed signal-to-noise ratios. *Audiology and Speech Research*, 13(1), 62-69.
- Kollmeier, B., Warzybok, A., Hochmuth, S., Zokoll, M. A., Uslar, V., Brand, T., et al. (2015). The multilingual matrix test: Principles, applications, and comparison across languages: A review. *International Journal of Audiology*, 54 Suppl 2, 3-16.
- Lee, J. H. & Yi, D. W. (2017). A comparison of adaptive sentence-in-noise tests. *Audiology and Speech Research*, 13(1), 9-18.
- Needleman, M. T., Wambacq, I., Besing, J., Spitzer, J. B., & Koehnke, J. (2015). The effect of background babble on working memory in young and middle-aged adults. *Journal of the American Academy of Audiology*, 26(3), 220-228.
- Nilsson, M., Soli, S. D., & Sullivan, J. A. (1994). Development of the hearing in noise test for the measurement of speech reception thresholds in quiet and in noise. *The Journal of the Acoustical Society of America*, 95(2), 1085-1099.
- Ozimek, E., Warzybok, A., & Kutzner, D. (2010). Polish sentence matrix test for speech intelligibility measurement in noise. *International Journal of Audiology*, 49(6), 444-454.
- Park, J. H. & Kwon, Y. C. (1989). Standardization of Korean version of Mini-Mental State Examination (MMSE-K) for use in the elderly. Part II. Diagnostic validity. *Journal of Korean Neuropsychiatric Association*, 28, 508-513.
- Polat, Z., Bulut, E., & Ataş, A. (2016). Assessment of the speech intelligibility performance of post lingual cochlear implant users at different signal-to-noise ratios using the Turkish Matrix Test. *Balkan Medical Journal*, 33(5), 532-538.
- Puglisi, G. E., Warzybok, A., Hochmuth, S., Visentin, C., Astolfi, A., Prodi, N., et al. (2015). An Italian matrix sentence test for the evaluation of speech intelligibility in noise. *International Journal of Audiology*, 54 Suppl 2, 44-50.
- Wardenga, N., Batsoulis, C., Wagener, K. C., Brand, T., Lenarz, T., & Maier, H. (2015). Do you hear the noise? The German matrix sentence test with a fixed noise level in subjects with normal hearing and hearing impairment. *International Journal of Audiology*, 54 Suppl 2, 71-79.
- Warzybok, A., Zokoll, M., Wardenga, N., Ozimek, E., Boboshko, M., & Kollmeier, B. (2015). Development of the Russian matrix sentence test. *International Journal of Audiology*, 54 Suppl 2, 35-43.
- Yi, D. (2016). Study on the construction and optimization of the Korean matrix sentence materials (unpublished master's thesis). Hallym University of Graduate Studies, Seoul.
- Zokoll, M. A., Fidan, D., Türkyılmaz, D., Hochmuth, S., Ergenç, İ., Sennaroglu, G., et al. (2015). Development and evaluation of the Turkish matrix sentence test. *International Journal of Audiology*, 54 Suppl 2, 51-61.