Korean Speech Dynamic Range According to Frequency, Gender, and Speech Type

Kyungju Lee¹, 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

주파수, 성별 및 어음 종류에 따른 한국어의 어음역동범위

이경주¹ · 허소희¹ · 진인기²

한림대학교 일반대학원 언어병리청각학과¹, 한림대학교 자연과학대학 언어청각학부 · 청각언어연구소²

Purpose: Dynamic range (DR) of speech refers to the minimum and maximum levels of speech across frequencies and it depends on language. Recent findings showed that Korean speech DR is different from the DRs of other languages such as English and Mandarin. The purpose of this study was to identify the characteristics of Korean speech DR according to frequency (21 bands), gender (male and female), and speech type [words, sentences, and continuous discourses (CDs)]. **Methods:** DRs for 70 participants were recorded and yielded using cumulative histogram levels. Then, the DRs were compared according to frequency bands, gender, and speech type after level normalization at 65 dB sound pressure level. **Results:** Although there were no statistical differences between DRs for sentences and CDs, DRs for words were wider than DRs for other speech types. In terms of gender, DRs for female speakers were wider than DRs for male speakers in all speech types except a few frequency bands. **Conclusion:** Speech DR is important in real ear measurements like speech mapping because audibility can be determined based on the relationship between a person's hearing thresholds and the DR of speech. Thus, the results of this study contribute to more accurate hearing aid fitting according to frequency bands, gender, and speech type for Korean listeners.

Key Words: Speech dynamic range, Speech intelligibility, Language, Hearing aid fitting.

Received: July 18, 2017 / **Revised:** August 22, 2017 / **Accepted:** September 11, 2017

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

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

INTRODUCTION

어음역동범위(speech dynamic range)는 주파수별 어음 에너지 분포 중 최대 레벨(maximum levels)과 최소 레벨(minimum levels) 사이를 의미하며 어음인지도 수행력(speech intelligibility performance)에 있어 중요한 역할을 한다(Jin et al., 2014; Moore, 2006). 예를 들어, 스피치 맵핑(speech mapping)과 같은 실이측정(real ear measurement) 시에 입력 음압에 따른 어음역동범위와 대상자의 청력역치 레벨 간의 관계를 통해서 보청기 적합의 결과를 예측할 수 있다(Moore, 2006). 만약 어음 인지도를 저해하는 요소(예: 소음 등)가 없다는 가정하에 한 사람의 청력역치 레벨이 모든 주파수 영역에서 어음의 역동범위

보다 좋다면, 100% 청취가 가능함을 의미한다. 반대로 한 사람의 청력역치 레벨이 모든 주파수 영역에서 어음의 역동범위보다 나쁘다면, 어떠한 말소리도 들을 수 없어 말소리 청취도가 0%임을 의미한다. 이렇듯 어음역동범위는 어음인지도 수행력의 결정에 중요한 역할을 한다.

어음역동범위는 다양한 요소에 의해 달라질 수 있다. 첫째, 어음역동범위는 주파수에 따라 다를 수 있다. Byrne et al.(1994)은 영어의 어음역동범위는 주파수별로 다른 범위를 가진다고 보고하였다. 예를 들어, 400 Hz의 어음역동범위는 4,000 Hz에 비해 약 7 dB 정도 넓은 어음역동범위를 갖는다고 보고하였다. Jin et al.(2014) 또한 영어를 비롯한 다른 언어에서도 주파수별로 다른 역동범위를 갖는다고 보고하였다. 한국어, 중국어(광동

어) 및 영어의 표준문장표에서 사용된 남성 화자들의 어음역동범위를 측정된 결과, 세 언어 모두에서 주파수별로 다른 역동범위가 관찰되었다. 예를 들어 한국어는 중심주파수(center frequency, CF)를 기준으로 4,800 Hz 영역에서 가장 넓은 역동범위(37 dB)를 보였고, 150 Hz 영역에서 가장 좁은 역동범위(13.75 dB)를 보였으며, 중국어의 경우에는 중심주파수를 기준으로 7,000 Hz 영역에서 가장 넓은 역동범위(38.75 dB)를 보였고, 150 Hz 영역에서 가장 좁은 역동범위(17.5 dB)를 갖는 것으로 나타났다.

어음역동범위는 언어에 따라서도 달라질 수 있다. Jin et al. (2014)은 한국어, 중국어(광둥어) 및 영어를 모국어로 사용하는 각 남성 성우의 목소리로 녹음된 음성 파일의 분석을 통해 언어별로 다른 어음역동범위를 가지고 있음을 보고하였다. 각 언어별 음성 샘플(sample)의 소리 레벨을 65 dB sound pressure level (SPL)로 정규화(normalization)한 후에 주파수별 어음역동범위를 측정된 결과, 세 언어 모두 다른 역동범위를 갖는 것으로 나타났다. 예를 들어, 한국어는 영어 및 중국어에 비해 500 Hz 이하에서 좁은 역동범위를 갖는 것으로 나타났고, 영어의 경우에는 중주파수(500~4,000 Hz) 영역에서 한국어와 중국어에 비해 좁은 역동범위를 갖는 것으로 나타났다. 하지만 4,000 Hz 이상의 고주파수 영역에서는 영어의 어음역동범위가 한국어 및 중국어에 비해 넓은 역동범위를 갖는 것으로 나타났다.

또한 어음역동범위는 성별에 따라서도 달라질 수 있다. Park & Jin(2015)은 남녀 화자의 목소리로 녹음된 한국표준문장표(Jang et al., 2008)의 어음역동범위를 분석한 결과, 성별에 따른 어음역동범위의 유의미한 차이가 주파수에 따라 나타날 수 있음을 보고하였다. 예를 들어, 중심주파수 기준 450 Hz 이하에서는 여성 화자의 어음역동범위가 넓은 것으로 나타났고, 450 Hz 이상에서는 남성 화자의 역동범위가 넓은 것으로 나타났다. 통계적으로는 450, 1,000, 1,170, 1,370, 1,600, 4,800, 7,000, 8,500 Hz(중심주파수) 영역에서 남녀 화자 간 어음역동범위의 차이가 유의미한 것으로 나타났다.

어음역동범위는 어음의 종류에 따라서도 달라질 수 있다. Boothroyd et al.(1994)은 영어 음소(phoneme)들은 주파수에 따라 다른 에너지 분포(level distribution)를 가진다고 보고하였다. 예를 들어, 자음 /m/은 중심주파수 기준으로 250 Hz 영역에서 가장 높은 에너지 분포도를 보였으며 역동범위는 약 15 dB 정도인 것으로 나타났다. 반면에 자음 /s/는 4,000 Hz 영역에서 가장 높은 에너지 분포도를 보였으며 역동범위는 약 10 dB 인 것으로 나타났다. Boothroyd et al.(1994)의 연구결과는 어음의 종류에 따른 어음역동범위의 차이를 직접적으로 보여주진 않았지만 어음을 구성하고 있는 음소들의 구성과 분포에 따라

어음역동범위의 특성이 주파수에 따라 달라질 수 있다는 점을 시사하고 있다.

위에서 언급한 선행연구들의 결과들을 요약하면 어음의 역동범위는 어음인지도 수행력에 있어 중요한 역할을 하며 주파수, 언어, 성별 및 어음의 종류에 따라 달라질 수 있는 가능성을 보여주고 있다(Byrne et al., 1994; Jin et al., 2014; Moore, 2006). 하지만 선행연구들의 대부분은 영어를 비롯한 외국어의 어음역동범위에 초점이 맞춰진 연구들이 주를 이루고 있으며(Boothroyd et al., 1994; Byrne et al., 1994), 한국어를 다룬 선행연구들의 경우에도 소수의 화자 데이터를 기반으로 연구가 이루어져 선행연구들에서 다뤄졌던 한국어의 어음역동범위 특성이 포괄적인 한국어의 어음역동범위로 다루기에는 한계가 있었다(Jin et al., 2014; Park & Jin, 2015). 이러한 한계점을 극복하기 위해서는 한국어의 어음역동범위를 다화자 음성 샘플을 기반으로 주파수, 성별 및 어음의 종류에 따라 산출하는 연구가 필요하다고 할 수 있다.

본 연구의 목적은 주파수, 성별 및 어음 종류에 따른 한국어의 어음역동범위를 측정하는 것이었다. 본 연구의 결과는 남녀에 따른 어음 종류별 어음역동범위의 특성을 주파수별로 제공할 것이다. 이는 어음인지지수(speech intelligibility index) 산출 및 실이측정 시에 한국어를 모국어로 사용하는 사람들을 위한 최적 적합(best-fit) 도출에 유용한 정보를 제공할 것이다.

MATERIALS AND METHODS

Participants

본 연구는 건청성인 남녀 각각 35명(총 70명)을 대상으로 하였다(평균 연령: 21.5세, 연령 분포: 20~28세). 본 연구에 참여한 대상자들의 순음청력검사 결과, 모든 참가자들은 250~8,000 Hz 영역에서 20 dB HL 내의 정상 범위를 보였으며 고막운동도 검사 결과는 A 타입으로 정상임을 확인하였다. 본 연구에 사용된 청력검사기기(GSI 61; Grason-Stadler, Eden Prairie, MN, USA)는 검사 시행 4개월 전에 연간 보정을 실시하였으며, 각 검사 시행 전 일일 보정을 실시하였다. 또한, 본 연구는 녹음을 통한 목소리의 분석이 주요한 요소이기 때문에 모든 대상자는 발성과 조음에 어려움이 없었으며 언어적 결함이 없는 대상자를 참여자로 선정하였다. 모든 연구 참여자에게 참여에 대한 수고비를 지급하였으며, 본 연구의 진행에 앞서 임상실험윤리위원회(Institute Review Board, IRB)의 승인을 받았다(승인번호: HIRB-2016-001). 또한 실험의 진행 이전에 모든 대상자들은 연구의 목적, 진행 절차 등에 대한 충분한 설명을 제공받았으며 각 참여자의 동의 후에 녹음을 진행하였다.

Recording and analysis procedure

본 연구는 방음실(doubled-wall)에서 진행되었으며, 음성분석프로그램[computerized speech lab (CSL), KayPENTAX™, Montvale, NJ, USA]과 마이크(Sennheiser e-835s, Wedemark, Germany)를 통해 녹음을 진행하였다. 녹음 시 표본화 주파수와 비트(bit)는 44,100 Hz와 16 비트로 설정하였다. 대상자는 편안한 자세로 의자에 앉은 상태에서 입은 마이크로부터 10 cm 거리를 유지하여 녹음을 진행하였다. 어음은 총 3가지(단음절, 문장 및 담화) 종류로 구분하여 녹음을 진행하였다. 단음절은 한국표준 일반용 단음절어표(Korean Standard Monosyllabic Word Lists for Adults)를 사용하였으며(Kim et al., 2008), 문장은 한국어 문장감사목록(Korean Speech Perception In Noise)을 사용하였다(Kim et al., 2000). 담화는 한국어로 전사된 동화인 ‘바람과 해님(Borim Press, Paju, Korea)’을 사용하였다. 녹음 과정에서 발음이나 조음이 부적합하다고 판단된 경우에는 재녹음을 실시하였으며, 피검자들이 녹음 도중 휴식을 원할 경우 수시로 휴식을 제공하였다.

녹음된 파일들을 대상으로 어음역동범위를 산출하기 위해 아래의 절차대로 분석을 진행하였다. 첫째, Adobe Audition CS6 (Adobe Systems Corp., San Jose, CA, USA)를 사용하여 모든 어음 종류(단음절, 문장, 담화)에 따른 녹음 파일들의 묵음(pauses)을 제거하였다. 묵음 시간이 200 ms 이상이 되면 어음의 발생 시에 나타나는 묵음의 범위에서 벗어나기 때문에 묵음 휴지(silent pauses)의 최소 시간을 250 ms로 설정하여 묵음을 제거하였다(Kendall, 2013). 다음으로 매트랩(Matlab, version R2017a; MathWorks Inc., Natick, MA, USA)을 이용하여 묵음이 제거된 데이터의 누적 히스토그램 레벨(cumulative histogram levels)을 도출하였다. 누적 히스토그램 레벨 도출 과정은 다음과 같다. 첫째, 녹음 데이터의 평균 실효값(root mean square) 레벨을 계산한 후, Praat v6.0.19 (Boersma & Weenink, 2016)을 통해 각 녹음 파일들을 65 dB SPL로 정규화하였다. 둘째, 중심주파수가 150~8,600 Hz의 범위를 가지는 21개의 임계 대역(critical bands)에서 신호 포락선(signal envelope)을 계산하였다. 이러한 21개의 주파수 대역 범위는 어음인지 지수를 계산할 때의 임계 대역 범위를 기준으로 하였다[American National Standards Institute (ANSI), 1997/R2012]. 셋째, 매트랩을 이용하여 1,000분의 1초 단위로 계산되는 상승 여현창(raised cosine window)을 통해 신호 포락선을 부드럽게 곡선화하였다. 넷째, 곡선화된 신호 포락선을 창 길이(window length)가 50% 중첩되게 하여 부표본화하고 dB SPL 단위로 변환시켰다. 마지막 단계로, 누적 신호 포락선 분포 레벨을 데시벨 신호 포락선 히스토그램으로 변화하여 어음의 역동범위를 도출하였다. 누적 히스토그램 레벨은 주파수 대역 내에서 어음의 누적

신호 포락선의 분포를 보여준다. 예를 들어, 99% 누적 히스토그램 레벨(L99)은 측정된 어음 신호에서 최고치 레벨(peak level)을 포함하거나 그 이하의 99% 부분을 가리킨다. 30% 누적 히스토그램 레벨(L30)은 측정된 어음 신호에서 최소 레벨을 포함하거나 그 이상의 30% 부분을 가리킨다. 다시 말하면, 본 연구에서 L99는 어음의 최대 레벨을 말하며, L30은 어음의 최소 레벨을 말한다. 어음역동범위는 최대 레벨에서 최소 레벨을 뺀 값을 나타내며(L99 - L30), 이 방식은 국제 어음 시험 신호(international speech test signal)를 개발하는 데 있어서 효율적인 어음 역동범위로 사용되었다(Holube et al., 2010).

Statistical analysis

본 연구는 녹음 데이터를 가지고 주파수별로 어음 종류와 성별에 따른 한국어의 어음역동범위를 비교하였다. 통계 분석은 SPSS-22 (IBM Corp., Armonk, NY, USA) 소프트웨어를 통해 분석하였고 일원분산분석(one-way ANOVA)을 이용하여 주

Figure 1. The dynamic range for Monosyllabic Words (square), Sentences (circle) and Continuous Discourses (triangle) as a function of 21 band frequencies for male (A) and female (B).

파수 대역에 따른 어음 종류별 분석을 실시하였다. 또한, 독립 표본 t 검정을 통해 어음별로 남녀 간의 차이를 분석하였다. 모든 통계 분석의 유의 수준은 0.05 미만으로, 사후 검정은 일원 분산분석을 시행한 경우에 한해서 Bonferroni test를 통하여 분석하였다.

RESULTS

어음 종류(단음절, 문장, 담화)에 따른 한국어의 어음역동범위를 Figure 1에 제시하였으며, 어음 종류 간의 통계적 유의성은 Table 1에 주파수별로 제시하였다. Figure 1에서 제시하듯이 단음절의 어음역동범위는 남성의 경우 7.79 dB (CF: 150 Hz)에서 27.71 dB (CF: 1,170 Hz), 여성의 경우 10.61 dB (CF: 150 Hz)에서 29.73 dB (CF: 1,370 Hz)의 범위를 가지는 것으로 나타났다. 문장의 어음역동범위는 남성의 경우 8.96 dB (CF: 150 Hz)에서 23.83 dB (CF: 7,000 Hz), 여성의 경우 11.06 dB (CF: 150 Hz)에서 23.19 dB (CF: 5,800 Hz)의 범위를 보였으며 담화의 어음역동범위는 남성이 8.64 dB (CF: 150 Hz)에서 24.66 dB (CF: 7,000 Hz), 여성이 11.61 dB (CF: 250 Hz)에서 23.52 dB (CF: 8,500 Hz)였다. Table 1에서 제시하듯이 단음절어의 어음역동범위가 남녀 모두 중주파수 대역에서 문장, 담화의 역동

범위보다 유의미하게 넓게 분석되었다. 남성의 경우, 중심주파수가 350~2,900 Hz일 때 단음절어의 역동범위가 문장, 담화의 역동범위보다 유의미하게 넓었다($p < 0.05$). 이러한 주파수 대역에서 단음절과 문장 간의 어음역동범위 차이는 1.36 dB (CF: 350 Hz)에서 7.16 dB (CF: 1,370 Hz) 사이였으며, 단음절과 담화 간의 어음역동범위 차이는 2.4 dB (CF: 350 Hz)에서 6.95 dB (CF: 1,170 Hz) 사이였다. 여성의 경우에는 중심주파수가 450~3,400 Hz일 때 단음절어의 역동범위가 문장, 담화의 역동범위보다 유의미하게 넓었다($p < 0.05$). 이러한 주파수 대역에서 단음절과 문장 간의 어음역동범위 차이는 2.23 dB (CF: 570 Hz)에서 7.6 dB (CF: 1,370 Hz) 사이였으며, 단음절과 담화 간의 어음역동범위 차이는 2.32 dB (CF: 3,400 Hz)에서 7.43 dB (CF: 1,370 Hz) 사이였다. Figure 1과 Table 1에서 볼 수 있듯이, 문장과 담화의 역동범위는 차이가 크지 않았다. 남성의 경우, 중심주파수가 350과 450 Hz일 때를 제외한 나머지 주파수 대역에서는 유의미한 차이가 나타나지 않았으며, 여성의 경우 중심주파수가 1,600과 1,850 Hz일 때를 제외하고 유의미한 차이가 나타나지 않았다.

Figure 2에서는 각 어음 종류별 최대 레벨(L99)과 최소 레벨(L30)을 통하여 어음역동범위(L99 - L30) 분포를 제시하였다. Figure 2에서 볼 수 있듯이 남녀 모두에서 대부분 단음절이 문

Table 1. Results of one-way ANOVA for dynamic ranges among different speech types across 21 frequency bands

Frequency (Hz)	MONO & sentences		MONO & CDs		Sentences & CDs	
	Male	Female	Male	Female	Male	Female
150		1.000	0.057	0.398	1.000	0.968
250			1.000	0.065	0.064	1.000
350		0.206				1.000
450						1.000
570					0.311	0.451
700					0.940	1.000
840					1.000	1.000
1,000					1.000	1.000
1,170					1.000	1.000
1,370					0.138	1.000
1,600					0.859	
1,850					0.536	
2,150					0.161	1.000
2,500					0.790	0.542
2,900					1.000	1.000
3,400	0.531				0.565	0.071
4,000	0.158	0.069	0.182		1.000	1.000
4,800		0.199	0.112		1.000	1.000
5,800	0.056		0.105		1.000	1.000
7,000	0.105		0.528		1.000	1.000
8,500	0.464	0.137	1.000	1.000	0.065	0.495

: Statistical significance ($p < 0.05$). MONO: monosyllabic words, CDs : continuous discourses

Figure 2. Maximum and minimum levels of dynamic range for monosyllabic words (straight line), sentences (dashed line) and continuous discourses (dotted line) for male (A) and female (B). SPL: sound pressure level.

장, 담화보다 더 큰 최대 레벨(L99)과 더 작은 최소 레벨(L30)의 분포를 가지는 것을 볼 수 있다. 남성의 경우, 단음절과 문장의 L99 차이값은 최대 3.99 dB (CF: 840 Hz)을 보였고 2 dB 이상 차이를 보이는 주파수 대역이 6개(CF: 700, 840, 1,000, 1,170, 1,370, 2,900 Hz)였으며, 단음절과 담화의 L99 차이값은 최대 4.13 dB (CF: 840 Hz)이었고 2 dB 이상 차이를 보이는 주파수 대역은 11개 영역이었다(CF: 350, 700, 840, 1,000, 1,170, 2,150, 2,500, 2,900, 3,400, 4,000, 4,800 Hz). 또한, 단음절과 문장의 L30 차이값은 최대 4.96 dB (CF: 1,370 Hz)이었고 2 dB 이상 차이를 보이는 주파수 대역이 8개(CF: 1,000, 1,170, 1,370, 1,600, 1,850, 4,800, 7,000, 8,500 Hz)였으며, 단음절과 담화의 L30 차이값은 최대 4.77 dB (CF: 1,370 Hz)이었고, 2 dB 이상 차이를 보이는 주파수 대역은 6개였다(CF: 570, 1,000, 1,170, 1,370, 1,600, 1,850 Hz). 여성의 경우, 단음절과 문장의 L99 차이값은 최대 3.11 dB (CF: 1,170 Hz)을 보였으며 2 dB 이상 차

이를 보이는 주파수 대역은 3개(CF: 1,000, 1,170, 1,370 Hz)였고, 단음절과 담화의 L99 차이값은 최대 2.7 dB (CF: 1,170 Hz)이었으며 2 dB 이상 차이를 보이는 주파수 대역은 2개였다(CF: 1,170, 5,800 Hz). 단음절과 문장의 L30 차이값은 최대 6.17 dB (CF: 1,600 Hz)이었고 2 dB 이상의 차이를 보이는 주파수 대역은 15개(CF: 700, 840, 1,000, 1,170, 1,370, 1,600, 1,850, 2,150, 2,500, 2,900, 3,400, 4,800, 5,800, 7,000, 8,500 Hz)였으며, 단음절과 담화의 L30 차이값은 최대 5.56 dB (CF: 1,370 Hz)이었고 2 dB 이상 차이를 보이는 주파수 대역은 11개였다(CF: 570, 700, 840, 1,000, 1,170, 1,370, 1,600, 1,850, 2,150, 2,500, 2,900 Hz). 반면, 남녀 모두에서 문장과 담화 사이의 L99, L30 값의 차이값은 상대적으로 크지 않았다. 남성의 경우, 문장과 담화 사이의 L99 차이값은 두 주파수 대역(CF: 3,400, 4,000 Hz)을 제외한 나머지 대역이 모두 2 dB 이내였으며 L30 차이값은 전 주파수 대역에서 2 dB 이내였다. 여성의 경우에도 문장과 담화 사이의 L99 차이값은 하나의 주파수 대역(CF: 4,800 Hz)을 제외한 나머지 대역이 모두 2 dB 이내였으며 L30 차이값은 전 주파수 대역에서 2 dB 이내로 나타났다.

어음역동범위의 남녀 비교는 각 어음(단음절, 문장, 담화)에 따라 각각 실시하였으며, 결과에 따른 그래프는 Figure 3에 제시하였고 남녀 간의 통계적 유의성은 Table 2에 주파수별로 제시하였다. Figure 3에서 볼 수 있듯이 세 가지 어음 모두에서 여성이 남성보다 전체적으로 더 넓은 어음역동범위를 가지는 것을 확인할 수 있었고, 3,700~5,300 Hz (CF: 4,000, 4,800 Hz)의 주파수 대역에서만 남성이 여성보다 유의미하게 더 넓은 어음역동범위를 가지는 것을 볼 수 있다. Table 2의 통계적 유의성을 같이 살펴보면, 단음절의 경우 중심주파수가 2,150 Hz (difference: 4.31 dB), 2,500 Hz (difference: 3.66 dB)를 포함한 11개의 주파수 대역에서 여성이 남성보다 어음역동범위가 유의미하게 넓었으며, 중심주파수가 4,000 Hz (difference: 3.36 dB), 4,800 Hz (difference: 7.51 dB)인 주파수 대역에서만 남성이 여성보다 유의미하게 더 넓은 어음역동범위를 가지고 있었다. 문장의 경우, 중심주파수가 1,370 Hz (difference: 2.6 dB), 2,500 Hz (difference: 2.83 dB)를 포함한 10개의 주파수 대역에서 여성이 남성보다 더 넓은 어음역동범위를 가지고 있었으며, 중심주파수가 3,400 Hz (difference: 1.37 dB), 4,000 Hz (difference: 2.93 dB), 4,800 Hz (difference: 5.26 dB)에서만 남성이 여성보다 유의미하게 더 넓은 어음역동범위를 가지고 있었다. 마지막으로, 담화의 경우 중심주파수가 1,000 Hz (difference: 3.36 dB), 2,150 Hz (difference: 3.47 dB)를 포함한 11개의 주파수 대역에서 여성이 남성보다 유의미하게 더 넓은 어음역동범위를 가지고 있었으며, 중심주파수가 4,000 Hz (difference: 3.14 dB), 4,800 Hz (difference: 7 dB)에서만 남성이

Table 2. Results of Independent samples t-test for dynamic ranges between male and female across 21 frequency bands

Frequency (Hz)	Monosyllabic words	Sentences	Continuous discourses
150			
250			
350	0.054		
450			
570	0.637	0.597	0.467
700	0.471	0.068	0.588
840			
1,000			
1,170		0.200	
1,370			
1,600			
1,850		0.077	
2,150			
2,500			0.580
2,900	0.821	0.577	0.421
3,400	0.209		0.854
4,000			
4,800			
5,800	0.402	0.442	0.578
7,000	0.920	0.269	0.133
8,500	0.316	0.476	0.683

: Statistical significance ($p < 0.05$)

여성보다 유의미하게 더 넓은 어음역동범위를 가지고 있었다.

Figure 4에서는 어음 종류별로 남성과 여성의 최대 레벨(L99)과 최소 레벨(L30)을 통해 성별 간의 어음역동범위(L99 - L30)를 제시하였다. 그림에서 볼 수 있듯이 전체적으로 여성이 남성보다 더 큰 최대 레벨과 최소 레벨을 가지고 있었다. 남녀 간의 차이를 살펴보았을 때 단음절의 경우, 남녀 간의 L99 차이값은 최대 4.61 dB (CF: 1,000 Hz)이었고 2 dB 이상 차이를 보이는 주파수 대역이 12개(CF: 250, 450, 570, 700, 840, 1,000, 1,170, 1,370, 1,600, 5,800, 7,000, 8,500 Hz)였으며, L30 차이값은 최대 7.34 dB (CF: 4,800 Hz)이었고 2 dB 이상 차이를 보이는 주파수 대역은 5개였다(CF: 570, 700, 840, 1,000, 4,800 Hz). 문장의 경우, 남녀 간의 L99 차이값은 최대 6.87 dB (CF: 840 Hz)이고 2 dB 이상의 차이를 보이는 주파수 대역은 13개(CF: 250, 570, 700, 840, 1,000, 1,170, 1,370, 1,600, 1,850, 2,900, 4,800, 5,800, 8,500 Hz)였으며, L30 차이값은 최대 4.49 dB (CF: 2,900 Hz)이며 2 dB 이상 차이를 보이는 주파수 대역은 10개였다(CF: 570, 700, 840, 1,000, 1,170, 2,900, 4,000, 4,800, 7,000, 8,500). 마지막으로, 담화의 경우 남녀 간의 L99 차이값은 최대 7.29 dB (CF: 840 Hz)이고 2 dB 이상 차이를 보이는 주파수 대역은 14개(CF: 250, 350, 450, 570, 700, 840, 1,000, 1,170, 1,370, 1,600, 1,850, 2,150, 2,900, 8,500 Hz)였으며, L30

차이값은 최대 7.42 dB (CF: 4,800 Hz)이고 2 dB 이상 차이를 보이는 주파수 대역은 9개였다(CF: 250, 570, 700, 840, 1,000, 2,900, 4,000, 4,800, 8,500 Hz).

DISCUSSIONS

본 연구에서는 주파수, 성별 및 어음 종류에 따른 한국어의 어음역동범위를 측정하고, 각 종류별 어음역동범위의 특징을 서로 비교해 보았다. 연구의 결과를 요약하면 다음과 같다. 첫째, 한국어의 어음역동범위는 어음에 따라 주파수별로 유의미한 차이가 있었다. 단음절의 경우에는 약 8~30 dB 정도의 어음역동범위를 보였으며, 문장의 경우에는 약 9~23 dB, 담화의 경우에는 9~25 dB 정도의 어음역동범위를 보였다. 어음에 따른 어음역동범위의 비교에서는 단음절의 어음역동범위가 문장 또는 담화의 어음역동범위보다 중주파수 영역에서 넓은 것으로 나타났다. 하지만 문장과 담화의 어음역동범위는 대부분의 주파수 대역에서 유의미한 차이가 나타나지 않았다. 둘째, 한국어의 어음역동범위는 성별에 따라 주파수별로 유의미한 차이가 있었다. 여성의 경우, 어음에 따라 어음의 역동범위의 차이가 있었지만 주파수에 따라 약 10~30 dB 정도의 어음역동범위를 보였고, 남성의 경우에는 약 7~27 dB 정도의 어음역동범위를 가지는 것

으로 나타났다. 성별에 따른 어음역동범위의 비교에서는 여성의 어음역동범위가 남성에 비해 대부분의 주파수 영역에서 넓은 것으로 나타났다. 어음에 따라 2개 내지 3개의 주파수 대역에

서만 남성의 어음역동범위가 여성에 비해 넓었으며 그 외는 유사하거나 여성의 어음역동범위가 남성에 비해 유의미하게 넓은 것으로 나타났다.

Figure 3. The dynamic range for male (square), female (circle) as a function of 21 band frequencies for monosyllabic words (A), sentences (B) and continuous discourses (C). *Statistical significance.

Figure 4. Maximum and minimum levels of dynamic range for male (straight line) and female (dashed line) for monosyllabic words (A), sentences (B) and continuous discourses (C). SPL: sound pressure level.

어음의 역동범위에서 최대 레벨의 분포는 어음인지도와 매우 밀접한 관계가 있다. Jin et al.(2017)은 어음의 역동범위가 어음 인지지수의 예측에 미치는 영향에 대해 보고하였는데, 어음의 역동범위가 넓어질수록 높은 어음인지도로 예측되는 이유는 어음인지지수의 계산식상 넓은 어음역동범위가 더 높은 최대 레벨을 포함하고 있어 신호대잡음비(signal-to-noise ratio)가 상승하기 때문임을 증명하였다. 본 연구의 결과들 중 최대 레벨의 관점에서 비교를 했을 때, 어음에서는 단음절이, 성별에서는 여성의 L99가 비교 군에 비해 유의미하게 높은 것으로 나타났다. 어음에 따른 비교에서는 맥락 단서(context cue)가 많은 문장 또는 담화의 경우, 실제 어음인지도가 단음절에 비해 높다는 선행연구의 결과(Depaolis, 1996)가 있기 때문에 최대 레벨의 분포가 어음인지도에 미치는 영향을 직접적으로 설명하기에는 한계가 있다. 하지만 성별에 따른 비교에서는 여성의 목소리가 남성의 목소리에 비해 넓은 역동범위를 가지며 특히 최대 레벨의 에너지가 남성에 비해 더 많이 분포하고 있기 때문에 동일 레벨에서 동일 어음을 들었을 때 더 높은 어음인지도 결과의 산출이 가능할 수 있음을 보여주고 있다. McBride et al.(2008)은 다양한 소음 레벨 상황에서 남성과 여성의 목소리로 어음인지도를 측정하였는데 소음의 레벨이 높아질수록 여성의 목소리로 청취했을 때 더 높은 어음인지도가 나타났다고 보고하였다. 이러한 선행 연구들의 결과를 고려했을 때, 한국어의 경우에도 평균적인 관점에서 여성의 목소리가 남성에 비해 더 좋은 어음인지를 산출할 수 있는 어음역동범위 특성을 가진다고 볼 수 있다.

영어의 경우에는 30 dB(평균 실험값 레벨 ± 15 dB)의 범위를 표준어음역동범위로 지정하여 어음인지지수 등에서 활용하고 있다(ANSI 1997/R2012). 영어의 경우에도 실제 측정된 주파수별 어음역동범위는 약 10~37 dB 사이로 다양한 분포를 보였지만(Boothroyd et al., 1994), 평균 실험값 레벨을 기준으로 +15 dB 정도의 최대 레벨을 포함하는 범위로 어음역동범위를 설정하면 어음인지지수에서 정확한 어음인지도 예측이 가능한 것으로 보고하고 있다(Beranek, 1947; Dunn & White, 1940; Pavlovic, 1988). 또한 Jin et al.(2017)은 한국어의 경우에도 최대 레벨을 포함하는 범위로 어음의 역동범위를 설정한다면 영어의 기준과 같은 30 dB를 적용했을 때 실제 측정된 어음인지도 점수와 어음인지지수에서 예측하는 어음인지도 점수와의 상관성이 매우 높게 나타난다고 보고하였다. 하지만 한국어의 어음역동범위를 다른 선행연구들의 경우에는 단일 화자에 근거한 역동범위만을 산출하고 사용하여 한국어의 표준어음역동범위를 산출하는 데 한계가 있었다(Jin et al., 2014; Jin et al., 2017). 이러한 한계를 극복하기 위해 본 연구에서는 다화자(70명)를 대상으로 어음역동범위를 산출하였고, 결과에서 제시한 L99 값은 어음 및 성별에 따른 한국어의 최적 어음역동범위 산출에 직접

적용될 수 있을 것이다.

실이측정의 방법 중 하나인 스피치 맵핑은 어음의 역동범위와 대상자의 청력역치 레벨을 비교하여 가청 영역에 해당하는 부분의 비율을 계산해 보청기 적합한 결과를 예측한다(Moore, 2006). 이때 적용되는 어음의 역동범위가 한국어를 기반으로 하고 단화자가 아닌 다화자의 평균값에 근거한 자료일 때, 한국어를 모국어로 사용하는 보청기 사용자를 위해 더욱 정확한 최적 적합(best-fit)을 가능하게 할 것이다. 또한 본 연구의 결과에서 알 수 있듯이 어음의 역동범위 특성은 어음, 성별에 따라 주파수별로 다를 수 있다. 이러한 특성 또한 보청기 적합 프로그램의 설정에서 어음 또는 화자에 따른 최적 적합 값을 구하는 데 있어 본 연구의 결과값이 기여할 수 있을 것이다.

중심 단어 : 어음역동범위 · 어음인지도 · 언어 · 보청기 적합.

Acknowledgments

이 성과는 2016년도 정부(미래창조과학부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임(NRF-2015R1C1A1A01052458).

REFERENCES

- ANSI. (1997). S3.5-1997, Methods for calculation of the speech intelligibility index. *American National Standards Institute*, 19, 90-119.
- Beranek, L. L. (1947). The design of speech communication systems. *Institute of Electrical and Electronics Engineers*, 35(9), 880-890.
- Boersma, P. & Weenink, D. (2016). Praat: Doing phonetics by computer. [Computer program]. Version 6.0.19.
- Boothroyd, A., Erickson, F. N., & Medwetsky, L. (1994). The hearing aid input: a phonemic approach to assessing the spectral distribution of speech. *Ear and Hearing*, 15(6), 432-442.
- Byrne, D., Dillon, H., Tran, K., Arlinger, S., Wilbraham, K., Cox, R., et al. (1994). An international comparison of long-term average speech spectra. *The Journal of the Acoustical Society of America*, 96(4), 2108-2120.
- 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.
- Dunn, H. K. & White, S. D. (1940). Statistical measurements on conversational speech. *The Journal of the Acoustical Society of America*, 11(3), 278-288.
- Holube, I., Fredelake, S., Vlaming, M., & Kollmeier, B. (2010). Development and analysis of an international speech test signal (ISTS). *International Journal of Audiology*, 49(12), 891-903.
- Jang, H. S., Lee, J. H., Lim, D. H., Lee, K. W., Jeon, A. R., & Jung, E. J. (2008). Development of Korean standard sentence lists for sentence recognition tests. *Audiology and Speech Research*, 4(2), 161-177.
- Jin, I. K., Kates, J. M., & Arehart, K. H. (2014). Dynamic range for speech materials in Korean, English, and Mandarin: a cross-language comparison. *Journal of Speech, Language, and Hearing Research*, 57(5), 2024-2030.
- 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.
- Kendall, T. (2013). *Speech Rate, Pause and Sociolinguistic Variation: Studies in Corpus Sociophonetics. (1st ed.)*. (pp. 51-79). Basingstoke: Palgrave Macmillan.
- Kim, J. S., Lim, D. H., Hong, H. N., Shin, H. W., Lee, K. D., Hong, B. N., et

- al. (2008). Development of Korean standard monosyllabic word lists for adults (KS-MWL-A). *Audiology and Speech Research*, 4(2), 126-140.
- Kim, J. S., Pae, S. Y., & Lee, J. H. (2000). Development of a test of Korean speech intelligibility in noise (KSPIN) using sentence materials with controlled word predictability. *Speech Science*, 7(6), 37-50.
- McBride, M., Hodges, M., & French, J. (2008). Speech intelligibility differences of male and female vocal signals transmitted through bone conduction in background noise: Implications for voice communication headset design. *International Journal of Industrial Ergonomics*, 38(11), 1038-1044.
- Moore, B. C. J. (2006). Speech mapping is a valuable tool for fitting and counseling patients. *The Hearing Journal*, 59(8), 26-28.
- Park, K. Y. & Jin, I. K. (2015). The dynamic range for Korean standard sentence material: A gender comparison in a male and a female speakers. *Journal of Audiology and Otology*, 19(2), 97-100.
- Pavlovic, C. V. (1988). Articulation index predictions of speech intelligibility in hearing aid selection. *ASHA*, 30(6-7), 63-65.