Speaker Gender Effects on Word Recognition

Angela Huang[1], Sara Jensen[1], Robert H. Margolis[1,2], Erica Williams[1], Ingrid McBride[1], Aparna Rao[1], George L. Saly[2], Richard H. Wilson[1]

[1] Department of Speech and Hearing Science, Arizona State University, Tempe, Arizona

[2] Audiology Incorporated, Arden Hills, Minnesota

Introduction

Speech recognition ability is routinely assessed in the clinical evaluation of hearing using open- and closed- set paradigms, typically with recorded word lists spoken by a male talker. Individuals with audiograms of similar severity and configuration can demonstrate varying abilities to extract meaning from speech and environmental stimuli. There is some evidence that recognition performance for speech materials recorded by a female talker are poorer than performance on materials recorded by a male talker[1,2]. Patients with sensorineural hearing loss report more difficulty understanding female speech. Comparisons are complicated by possible effects of different calibration methods and presentation levels and other acoustic differences in the speech of male and female talkers. A comparison of recognition performance for words recorded by male and female talkers that are carefully matched for level has not been reported. In this study we measured word recognition performance in normal hearing and hearing-impaired listeners for monosyllabic CNC words spoken by a male and a female talker.

Methods

Speech Materials

Monosyllabic words recorded by a female talker (VA recordings[3]) and a male talker (Auditec recordings[4]) were digitally adjusted so the rms level in a 50-ms interval in the central portion of the vowel was identical for each word. Additionally, the overall peak amplitudes of each monosyllabic word for each recording were measured using Adobe Audacity (Figure 1).

Figure 1. Concatenated waveforms of the 25-word lists used in this study. The top trace is the VA (female) recording. The Auditec (male) recording is below.

Subjects:

One ear of ten normal-hearing subjects was tested. Thresholds at octave frequencies (250 – 8000 Hz) did not exceed 25 dB HL. The average 3-frequency pure tone average (500, 1000, 2000 Hz) was 4.8 dB HL. Subjects were screened with otoscopy and tympanometry and reported no recent signs of otologic disease.

One ear of five subjects with SNHL was tested. The subjects had normal otoscopy and tympanograms and had no recent history of middle ear disease. The audiograms of the test ear of the subjects with SNHL are shown in Figure 2.

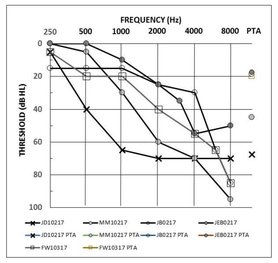

Figure 2. Audiometric thresholds and 3-frequency pure-tone averages for the five SNHL subjects.

Procedures

Speech recognition was tested using recorded monosyllabic words in a four interval forced-choice paradigm. The test word and three rhyming foils were presented on a touch screen. The subjects touched the word they thought they heard. 25 words were presented at five levels (5 - 25 dB re: pure tone average in 5 dB steps). SNHL subjects were tested at two additional levels (up to 35 dB SL).

Results

Normal-Hearing Subjects

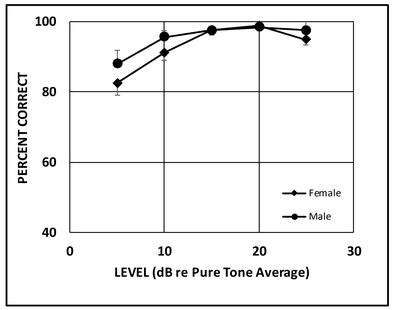

Figure 3. Psychometric functions for the VA (female) recording and Auditec (male) recording in normal hearing subjects.

Closed-set word-recognition scores were lower for the female speech than for male speech. The performance-intensity function for female speech is shifted toward higher levels by about 4 dB. Error bars are the standard error of the mean.

SNHL Subjects

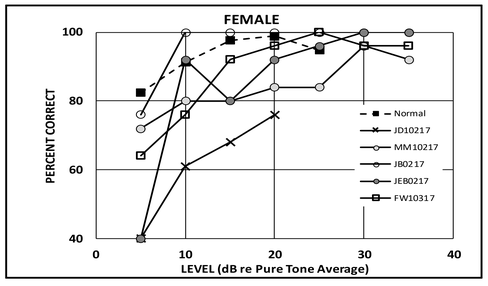

Figure 4. Psychometric functions for the VA (female) recording in SNHL subjects.

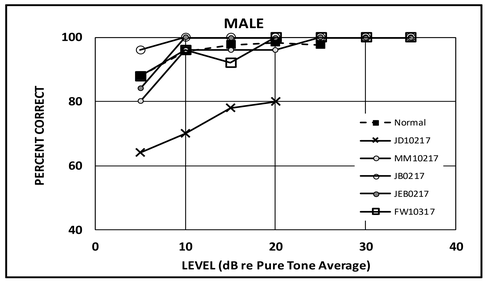

Figure 5. Psychometric functions for Auditec (male) recording in SNHL subjects.

Subjects with SNHL had higher recognition scores for male speech than for female speech at pre-asymptotic levels.

Speech Analysis

	Crest Factor (dB)		Peak Location (ms)	
	Mean	S.D.	Mean	S.D.
Female	6.93	0.87	0.12	0.05
Male	10.28	1.47	0.16	0.07
Diff (M-F)	**3.34**		**-0.04**	

Table 1. Mean crest factors (difference between the peak amplitude and the rms level of the vowel in dB) and mean peak locations in ms for the female and male recordings.

A difference of 3.34 dB in mean crest factor was obtained between the male and female recordings indicating the male recordings to have an average of 3.34 dB peak amplitude greater than female recordings.

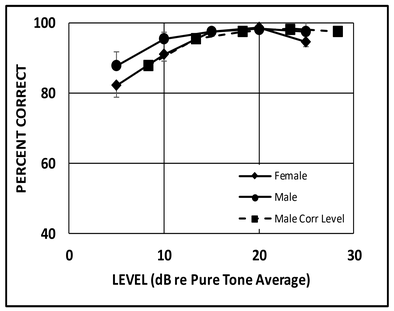

Figure 6. Psychometric functions for the VA (female) recording and Auditec (male) recording in normal hearing subjects in addition to a corrected performance-intensity curve for the male recording using the 3.34 crest factor difference calculated in Table 1.

Shifting the performance-intensity function to equalize crest factor for female and male speech eliminated the gender effect on speech recognition difference.

Conclusions

Forward Speech Recognition

1. In a closed set paradigm, normal-hearing listeners and listeners with SNHL have higher recognition scores for male speech than for female speech when word levels are adjusted to equalize the level in the steady state portion of the vowel of monosyllabic CNC words. For normal subjects the difference is equivalent to a shift of approximately 4-dB in the psychometric function.

2. The scores of subjects with SNHL were lower than those for normal-hearing subjects. Most approached 100% at the highest levels. In a previous study with a larger sample of subjects with SNHL some listeners achieved significantly reduced scores at high levels[2].

3. The levels of the words were matched based on the rms level in a 50-ms Interval in the central vowel. It is possible that another method for matching the levels of the words would produce equivalent performance.

4. The mean difference of 3.34 dB in crest factor between the male and female recordings accounted for the performance differences seen at low presentations levels. Discrepancies remain in high levels.

5. A larger sample size is needed to understand the performance differences between male and female recordings at high levels.

References

[1]Wilson, RH, Zizz, CA, & Shanks, JE. (1990). Normative Data in Quiet, Broadband Noise, and Competing Message for Northwestern University Auditory Test No. 6 by a Female Speaker. *Journal of Speech and Hearing Disorders; 55*:771-778.

[2]Margolis, RH, Gilbert, HM, Madsen, BM, Wilson, RH, Saly, GL. Automated Forced Choice Word-Recognition Test. Presented to the American Auditory Society, March 3, 2016. Available at http://audiologyincorporated.com/articles.

[3]Department of Veterans Affairs. Speech Recognition and Identification Materials, Disc 4.0. Produced by Auditory Research Laboratory, VA Medical Center, Mountain Home, Tennessee.

[4]Auditec, Inc. NU-6 (Northwestern University Auditory Test #6). Available from www.auditec.com.

This poster is available at http://audiologyincorporated.com/articles

For further information, please contact Robert H. Margolis (rhmargo001@gmail.com).