SPEECH PERCEPTION TESTING WITH REVERSED SPEECH

Roselynn R. Schroeder[1], Aparna Rao[1], Robert H. Margolis[1,2], Bryce Altus[1]
[1] Department of Speech and Hearing Science, Arizona State University, Tempe, Arizona
[2] Audiology Incorporated, Arden Hills, Minnesota

Introduction

Damage to the cochlea can lead to impaired speech communication in individuals with sensorineural hearing loss (SNHL). Clinical speech recognition tests aim to evaluate the speech perception deficit. However the development, validation, appropriate scoring, and accurate translation of these tests in every language is a daunting task. Access to diagnostic audiology is currently constrained by the limited range of developed and validated word recognition tests. This project aims to explore a method that may be capable of testing speech perception with speech-like, non-linguistic stimuli.

Speech perception involves processes from spectro-temporal analysis to lexical representation (Poeppel et al. 2008; Pittman et al., 2017). The first stage of speech processing is the spectro-temporal analysis in the cochlea. Lexical representation then aids in word identification. When speech is reversed, the long-term spectral content is preserved while temporal features are disrupted (Sheffert et al 2002). Identification will then be dependent solely on the comparison of spectro-temporal properties to those of the target.

We hypothesized that participants with hearing loss would perform worse than those with normal hearing in the reversed speech condition due to degradation of spectro-temporal analysis.

Methods

Speech Materials

Monosyllabic words recorded by a female talker (VA recordings[1]) and a male talker (Auditec recordings[2]) were digitally adjusted so the rms level in a 50-ms interval in the central portion of the vowel was identical for each word. Backward speech stimuli were the same words, digitally reversed with sound processing software (Adobe Audition).

Participants

- All participants were screened with otoscopy and tympanometry and reported no recent signs of otologic disease.
- One ear of 14 normal-hearing participants was tested.
  - Average age= 26.1 years, range: 22-30 years, 8 males ; 6 females.
  - 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 5.6 dB HL.
- One ear of 12 participants with sensorineural hearing loss was tested.
  - Average age= 69.2 years, range: 57-78 years, 8 males; 4 females.
  - The average 3-frequency pure tone average (500, 1000,2000 Hz) was 26.2 dB HL. Only two pure tone averages exceeded 35 dB HL. Average audiograms are shown in Figure 2.

Average thresholds for participants with sensorineural hearing loss

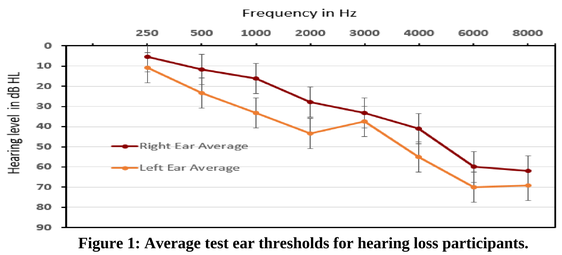

Figure 1: Average test ear thresholds for hearing loss participants.

Forward Speech Recognition

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

Backward Speech Recognition

The same monosyllabic word recordings were digitally reversed and presented in a three-interval, forced-choice paradigm. A target "word" was presented followed by three tokens, one of which was identical to the target, and the other two were randomly selected foils. The listener touched a button on the touchscreen (labeled 1, 2, or 3) to indicate which of the tokens matched the target. Presentation levels were the same as those used for the forward speech experiment

Results

Speech Mode vs Level Interactions

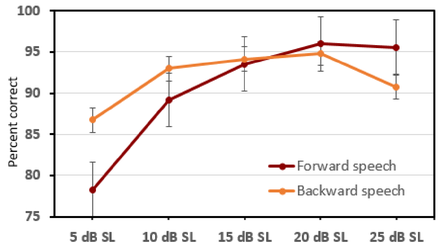

Figure 2: Speech mode × level ($p < .01$) Performance with forward speech was more difficult at low SLs. Performance improved at high SLs with scores greater than those obtained with backwards speech at levels above 15 dB SL.

Speech Mode vs. Speaker Gender

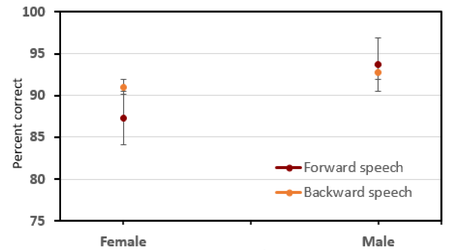

Figure 3: speech mode × speaker gender ($p < .02$). Performance was better with female talker in the backward speech mode, whereas it was comparable for male talkers in the two modes.

Normal hearing Participants: Performance-Intensity functions

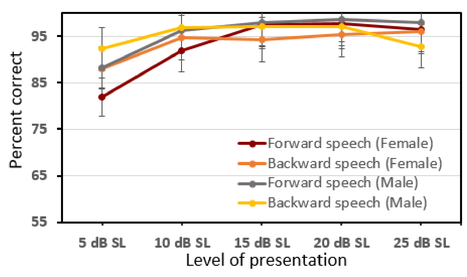

Figure 4: Performance-Intensity functions obtained with normal hearing Participants with the two modes of presentation. Scores were not significantly different for speech mode × level × gender.

SNHL Participants: Performance-Intensity functions

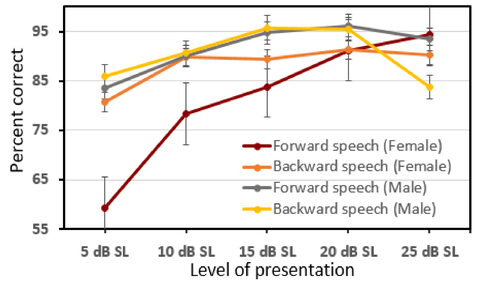

Figure 5: Performance-Intensity functions obtained with normal hearing Participants with the two modes of presentation. Scores were not significantly different for speech mode × level × gender.

Conclusions

- Recognition of backward monosyllabic words from one another was an easy task for normal-hearing and participants with sensorineural hearing loss.
- In a closed-set task, word recognition with forward speech was difficult at low SLs possibly due to the lexical processing load associated with the task.
- Lexical processing load could have contributed to the gender effects seen. Word recognition performance for forward speech with a female talker was significantly poorer at low SLs of presentation compared with performance in all other conditions. This pattern was consistent across groups.
- Overall performance was similar for the two groups possibly due to the following reasons:
  - The forced-choice response paradigm using three different reversed words was an easy task even for participants with hearing loss.
  - The participants chosen for the experiments were mostly limited to mild to borderline moderate loss due to the limits of the equipment. Participants with greater degrees of hearing loss may show different results.
- Lists used had only 25 words. Variability in performance is higher with a shorter list (compared with 50 word lists).

Future Directions

Low presentation levels were necessary to avoid scores that were too high to discriminate between groups. A task that stresses the spectro-temporal analysis capability of the listener may be more sensitive to the effects of sensorineural hearing loss. This would permit testing at higher stimulus levels that are more closely related to typical listening levels.

References

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

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

Pittman, A. L., Stewart, E. C., Odgear, I. S., & Willman, A. P. (2017). Detecting and learning new words: The impact of advancing age and hearing loss. *American Journal of Audiology*, *26*(3), 318-327.

Poeppel, D., Idsardi, W., & Van Wassenhove, V. (2008). Speech perception at the interface of ne urobiology and linguistics. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *363*(1493), 1071-1086.

Sheffert, Sonya M., Pisoni, David B., Fellowes, Jennifer M., & Remez, Robert E. (2002). Learning to Recognize Talkers from Natural, Sinewave, and Reversed Speech Samples. *Journal of Experimental Psychology: Human Perception and Performance*,*28*(6), 1447-69.

This poster is available at http://audiologyincorporated.com/articles
For further information, please contact Robert H. Margolis (rhmargo001@gmail.com).