# GENDER EFFECT IN A NONSENSE SYLLABLE TEST

Petri Korhonen, M.Sc. (Tech.),
Widex Office of Research in Clinical Amplification (ORCA), Widex USA

## ABSTRACT

Effect of speaker gender in bisyllable nonsense open-set speech identification test was studied. Four female and four male speakers were recruited to produce the source material. One female and male voice were selected for a final speech test implementation based on clarity and correctness of their pronunciation. Acoustic analysis was carried out to evaluate the differences in the acoustic cues produced by the two speakers. In addition, 9 normal hearing test subjects were recruited to evaluate the phoneme identification with five different low-pass filtering conditions. Objective and subjective measures providing details of gender effects in speech test development are reported.

## INTRODUCTION

A nonsense speech test was developed to allow systematic analysis of phoneme level error patterns. Typically speech tests have been implemented using only one speaker; more specifically a male voice. The frequency characteristics of speech produced by female and male speakers are different. Higher fundamental frequency, typical for female speakers, creates more sparse harmonic structures for voiced phonemes. Also, high frequency cues of fricative sounds produced by female speaker occupy a higher frequency region than those produced by a male speaker. The sensitivity of the speech test may be compromised if male speaker is used for testing the identification of fricative sounds, which have the energy concentrated at the higher frequencies. In addition, this difference has a potential of resulting in different phoneme level error-patterns and should be addressed in interpretation of test results.

To study the gender differences, two separate implementations of the test were produced, one with source material spoken by a female talker and one with one male talker. The current study evaluated the performance differences on these two test implementations using objective and subjective measures.

The speech test used in this study was designed specifically as an analysis tool for testing the effectiveness of hearing aid features which primary goal is to improve the audibility of higher frequencies. The new speech test stimuli consisted of 25 American English consonants coupled with five vowels in initial, medial, and final positions resulting in as list of 115 CVCVC nonsense syllables. Source material was recorded using 44.1kHz audio sampling frequency in an audiometric test booth using a high-fidelity condenser microphone. The recorded source material was edited and equalized for uniform presentation level using 50 ms sliding window peak RMS.

## METHODS

### *Acoustic analysis*

Spectral analysis was carried out to identify the spectral ranges and spectral centroids for all consonants. The locations of spectral peaks and shapes of spectral envelopes were measured using FFT. The analyis was carried out on each phoneme, and phoneme position individually.

### *Identification test*

Nine normal hearing subjects (determined as HL less than 10dB at 500Hz, 1000Hz, 2000Hz, 4000Hz) were recruited to evaluate the effects of gender in the identification of the source material. In addition to using wideband stimuli, low-pass filtering conditions with cut-off frequencies 4000Hz, 2000Hz, 1500Hz, 1000Hz, and 500Hz were applied to the source material to systematically reduce the availability of high frequency cues. Differrent test conditions used in different trials were counter balanced.

## RESULTS

### *A. All phonemes*

- Measured long term spectrum was similar between the two speakers.
- The female speaker has around 4dB lower spectrum at 2300-3300Hz; 5dB lower at 5000-6800Hz; and 5dB higher at 11200-13000Hz.
- The difference in identification scores between the male and female speakers was between 1.5-7.9% depending on the filtering cut-off frequency ($p<0.05$ at 2000Hz, 1500Hz, and 500Hz) (Figure 1).

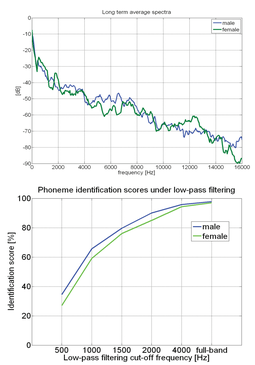



*Figure 1. The long term acoustic spectrum and identification results for all phonemes*

## RESULTS (CONT.)

### *B. Post-alveolar fricatives*

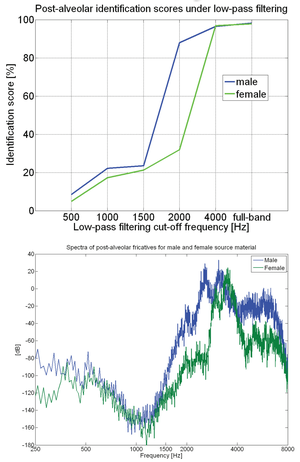



*Figure 2. Identification of post-alveolar fricatives and their averaged spectra*

An example of the phonemes that demonstarte the gender differences effectively in both acoustic analysis and in identification test results is post-alveolar fricatives (/ʃ/, and /ʒ/)). The identification score at 2000Hz is 55% lower for source material produced by a female speaker than for material produced by a male speech material. Acoustic analysis also shows that the location of the spectral peak of the female speaker is around 1000Hz lower than that of the male speaker (2000Hz vs. 3300Hz).

### *C. Fricatives*

- The acoustic analysis for fricative /s/ showed the differences in the location of the spectral peaks (male: 4000-8000Hz; female: 5600Hz-9600Hz).
- The level of the frication peak is approximately 30dB lower for this particular speaker.
- The low-pass filtering conditions used in the current study were bellow both of these peak frequencies.

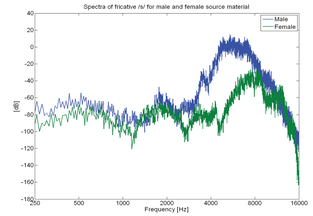
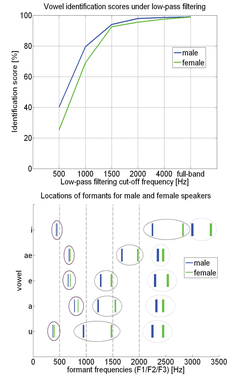



*Figure 3. The acoustic spectra of voiceless fricative /s/*

## RESULTS (CONT.)

### *D. Vowels*

- Effects of speaker gender differences are not unique to phonemes with high-frequency componets caused by frication.
- Differences in the formant locations between male and female speakers are a result of the differences in vocal tract lengths.




- Measured formant peak locations for the materials used in the current test corresponded closely to those reported in literature for average male and female speakers
- The greatest differences between the male and female formants were measured at F2 where the difference was up to 600Hz (/i/).
- The identification accuracy was 9.2% lower for female speech material with 1000Hz cut-off frequency ($p=0.044$); and 15.1% with 500Hz cut-off frequency ($p = 0.001$).

*Figure 4: Identification scores for vowels and locations of the first three formants F1, F2, and F3*

## DISCUSSION

The differences in nonsense speech test material produced by a male and a female speaker were studied using acoustic analysis and listening tests. The differences in the acoustic characteristics and in identification performance between the source material produced by the male and female speakers were more pronounced for only some speech sounds. The effect was greatest in identification of post-alveolar consonants. Other phoneme groups affected by the gender differences were vowels, unvoiced fricatives, nasals, and approximants.

When evaluating the effectiveness of hearing aid features designed to address the audibility of high frequency cues, the effects may be only seen for a particular set of speech sounds. In addition, these effects may be more readily seen for sounds produced by a female speaker than by a male. A careful analysis of the source material at phoneme level helps in the correct interpretation of the results, and ensures that those effects which are specific to only subset of speech sounds will be noticed.

## REFERENCES

Stelmachowicz PG, Pittman AL, Hoover BM, Lewis DE. (2001) Effect of stimulus band width on the perception of /s/ in normal- and hearing-impaired children and adults, The Journal of the Acoustical Society of America 110(4): 2183-2190.