



Reception Thresholds for Sentences: Female Talker Materials

Michael Nilsson

Rebecca Nilsson

Victor Bray

ABSTRACT

Objective measures of speech recognition in quiet and noise are becoming the standard with which rehabilitation and hearing aid signal processing effectiveness are judged. Extensive research protocols require larger sets of materials to eliminate learning effects, as well as alternative voices to increase the generalizability of findings. The development of a new set of materials following the methodology of the Hearing In Noise Test¹ will be presented, but using new sentences and a female voice.

INTRODUCTION

The measurement of speech recognition performance using sentence materials has become a benchmark for the behavioral evaluation of the effectiveness of various hearing aid signal processing strategies. English language materials have been available for many years, but only recently have sets of equivalent materials been in the public domain. As this type of testing gains popularity, the demand for additional materials increases. Creating additional materials also minimizes long-term learning effects from repeated testing with the same materials. This presentation will review the development to-date of a new set of sentence materials for use in measuring adaptive thresholds in quiet and noise that is intended to supplement the materials already available. The materials include all new sentences spoken by two different female talkers. Development is still in progress, and normative procedures to equate sentence difficulty will be the next step in completing the materials.

METHODS

Sentence Development

A large sample of short (5-7 syllable), easy (derived from 2nd grade reading materials) sentences were collected for use in the recordings. A large sample is necessary, as recordings will be dropped that cannot be scaled for equal difficulty.

A total of 525 sentences were collected that were deemed "conversational" in their style, simple in their grammatical structure, and sufficiently different from the HINT sentences to prevent learning effects from transferring over from one set of materials to another.

Recordings

Two alternate female voices were chosen, one with a softer, higher pitched voice, and the other with a louder, lower pitched voice. Both talkers spent about 2 hours reading sentences, and repeating any recordings deemed unclear (extraneous sounds, stumbles or flubs, word changes, etc.).

Recordings were made in a double-walled audiometric test room with the talker 1 meter from a B&K 4144 1-inch microphone. Signals were routed through a B&K 2260 Investigator which monitored levels for clipping, and amplified the signal going into a Creative Labs soundcard. Signals were recorded into digital files using the soundcard and Soundforge software. Recordings were made to mono files with 16-bit resolution and a 44.1 kHz sampling rates.

Groups of sentences were recorded to prevent the need to continuously start and stop the software.

RESULTS

Editing

Following the recording, all sentences were edited into individual files without any extra silence before or after the sentence. Individual RMS levels, as well as maximum and minimum digital values were recorded.

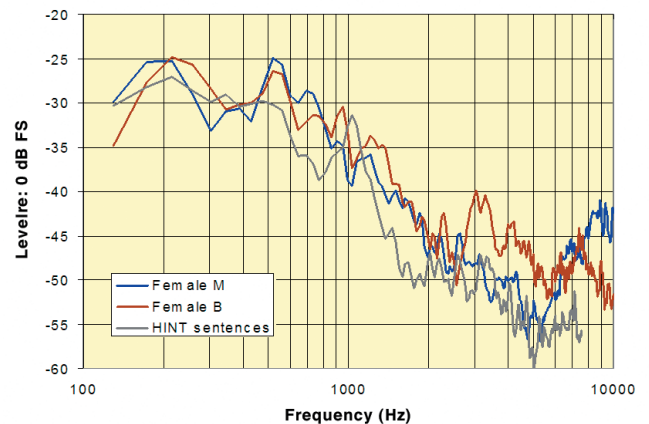


Figure 1: Calculated spectra using a 1024-point FFT and Blackman-Harris windowing for one minute samples of the two Female Voices and a 17-second sample of the original HINT recordings.

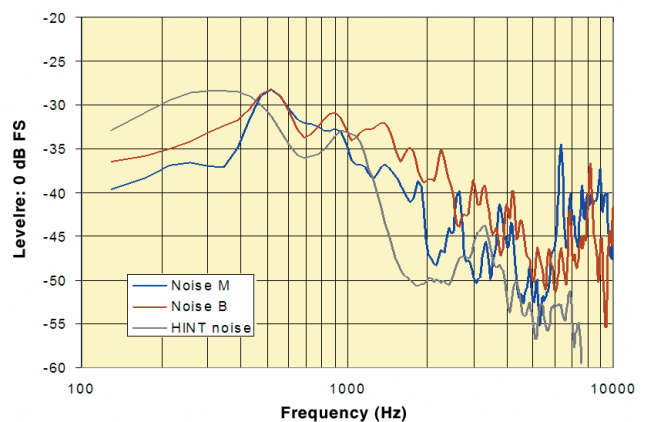


Figure 2: Calculated spectra using a 1024-point FFT and Blackman-Harris windowing for three minute samples of the two new noises and a 17-second sample of the original HINT noise.

Scaling

All sentences were scaled to a common RMS value that was chosen to avoid any digital clipping while providing maximum use of the digital space. The level was chosen to be the same for both female voices.

Calculation of Long-Term Spectrum

A waveform file consisting of concatenated sentences was created to calculate a long-term spectrum for each female voice. Cool Edit Pro, Version 2, software was used, as the calculated spectrum could be saved as numeric values in

Excel. The calculated spectrum for each female voice, as well as the spectrum of the original HINT sentences, are displayed in Figure 1. Figure 2 shows the first pass at spectrally matched noise for the female voices, as well as the original HINT noise.

CONCLUSIONS

A comparison of the spectra of the various materials shows the expected differences between the original male and new female voices, as well as the smoothed response of the original HINT noise. Differences between the two female voices can also be seen, including a difference in the fundamental frequency associated with the perceived difference in pitch between the two.

It is also important to note the high frequency differences. The original HINT recordings were made at a sampling rate of 20 kHz, and were low-pass filtered at 8 kHz to prevent aliasing. The new recordings do not have this limitation, and therefore retain much more high-frequency information.

The next step in the test development will be to complete the creation of a spectrally matched noise for each female talker, and present the sentences in noise at a fixed SNR to compare individual difficulty of each sentence. The goal is to end up with approximately 360 sentences of equal difficulty from each talker that can be used to measure adaptive thresholds.

REFERENCES

1. Nilsson, M.J., Soli., S.D., and Sullivan, J. (1994). Development of the Hearing In Noise Test for the measurement of speech reception thresholds in quiet and noise. *J. Acoust. Soc. Am.*, 95(2), 1085-1099.

