Vol. 5 No. 3 / 2002

SONIC innovations

Test Paradigm Manipulation During the Evaluation of Speech Recognition in Noise

Michael Nilsson

Robert M. Ghent

Victor Bray

Abstract

Benefit from amplification can be expressed in many domains, one of which is speech recognition, either in quiet or noise. One method in wide-spread use, is the Hearing In Noise Test (HINT1) which measures a Reception Threshold for Sentences (RTS) using an adaptive technique. Originally designed with a single, moveable source and a fixed presentation time, several modifications to the materials have been investigated to understand the utility and generalization of HINT performance to other, non-clinical conditions, as well as its sensitivity to testing advanced signal processing features available in new hearing aids. Noise onset time, single versus multiple noise sources, correlated versus uncorrelated maskers, and steady-state versus modulated maskers (multi-talker speech) were all manipulated in a multi-site evaluation using normal-hearing listeners. Data presented indicate that performance is highly dependent upon the level of these manipulations, necessitating careful selection and control of test conditions when evaluating benefit from signal processing.

International Hearing Aid Research Conference August 21-25, 2002 Tahoe City, California

INTRODUCTION

There are many competing requirements that must be considered when designing test methods to assess human performance with hearing aids. When quantifying speech intelligibility in noise (especially with advanced hearing aids), temporal characteristics (used by noise reduction algorithms), and spatial characteristics (used by directional microphone systems) must be controlled or consistently manipulated to allow for parsing of individual contributors to performance.

The Hearing In Noise Test (HINT) was used as the starting point1. The HINT is a pre-recorded test used to measure Reception Thresholds for Sentences (RTS) in quiet or in noise. The temporal and spectral characteristics of the original masking noise are stationary, and the noise spectrum matches the long-term average spectrum of the sentences. Several characteristics were evaluated, including noise onset time, and the ability to measure RTS with multi-talker maskers.

The test environment was modified to compare the single masker condition to the presentation of maskers in a quasi-free sound field as defined in Section 5.3 of ISO 8253-22. For a directional hearing device to show benefit, spatial separation of target and masker is required. This allows some directional advantage and simulates situations typically described as problematic for hearing impaired listeners. The test setup was replicated at three sites and normative data were collected at all sites to evaluate the equivalency of the test setups and to allow comparisons of data collected from samples of different populations.

METHODS

Test Materials

Several modifications were made to the HINT masking noise for the present study. The first modification was to move the start of the noise forward in time and increase the time between the onset of the noise and the onset of the target sentences from 0.5 seconds to approximately 5.0 seconds. The second modification to the HINT noise involved the creation of four correlated (time-coherent) and four uncorrelated Multi-Talker Speech (MTS) masker tracks to be reproduced simultaneously from four different loudspeakers arranged in a sound field array (see Figure 1). The third modification was the creation of a set of uncorrelated maskers with four, eight, twelve, and sixteen individual talkers reading the "Television Passage" 3.

Test Equipment

The system was calibrated such that the HINT masking noise was reproduced at 65 dB(A) at the center of the listening position, under any test condition (masker from the front loudspeaker, or masker from the sound field array). When the sound field array was used to reproduce the masking condition (both HINT noise and MTS), the settings ensured equal contribution from each loudspeaker. Replication at the alternate test sites was accomplished using identical equipment and bit-for-bit copies of the digitally recorded test materials. After connecting the equipment, calibration at the listening position was verified to be within 0.5 dB of the original test site.

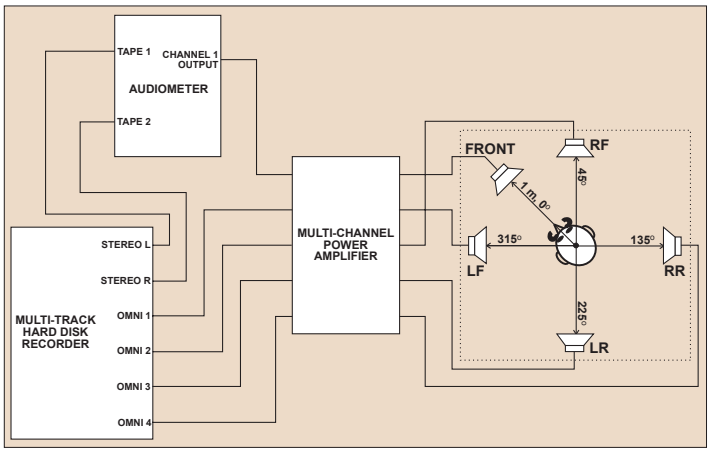

Figure 1: Equipment diagram and soundfield layout

Subjects

Subjects included 10 normal-hearing, native speakers of English with pure-tone thresholds less than 20 dB HL between 250 and 6000 Hz tested at sites 1 and 2, and 20 subjects at site 3.

RESULTS

Alternate Noise-Onset Times

RTS were measured using 20-sentence lists with both 0.5-second noise onset and with 5.0-second noise onset. An analysis of variance (ANOVA) found a significant difference in RTS for the two conditions [$F(1, 9) = 38.55$, $p < .01$]. The longer noise onset produced lower RTS (-8.16 dB SNR) than the short noise onset (-5.58 dB SNR).

Correlated versus Uncorrelated Noise Sources

RTS were measured using a single noise file played through all four loudspeakers simultaneously (a coherent noise field condition) versus four separate tracks of uncorrelated noise played through individual loudspeakers surrounding the listener (a non-coherent noise field condition). An ANOVA found significant differences between the correlated and uncorrelated conditions [$F(1, 9) = 6.28$, $p < .05$], with lower RTS in the correlated noise condition (-8.16 dB SNR) than in the uncorrelated noise condition (-7.27 dB SNR).

Multi-talker Speech Masker at Multiple Sites

RTS were measured on normal-hearing subjects in six conditions at three test sites. Of interest were differences in absolute performance with various maskers. All conditions used a noise-onset time of 5.0 seconds, and the diffuse field conditions all used uncorrelated signals.

An ANOVA was used to compare test site and masker type (HINT noise presented from the front loudspeaker, uncorrelated HINT noise presented from four loudspeakers, and one, two, three, or four individual talkers played out of each of the four sound field loudspeakers creating 4-, 8-, 12-, or 16-talker speech maskers). A main effect of test site [$F(2, 37) = 29.93$, $p < .01$] was found, with a post-hoc analysis revealing that site 1 was different from the other two sites (see Figure 2). A main effect of type of noise was found [$F(5, 185) = 43.67$, $p < .01$], with post-hoc analyses showing the noise front, 8-, and 12-talker speech were not different, nor were the diffuse noise and 4-talker speech, but all other comparisons were significant.

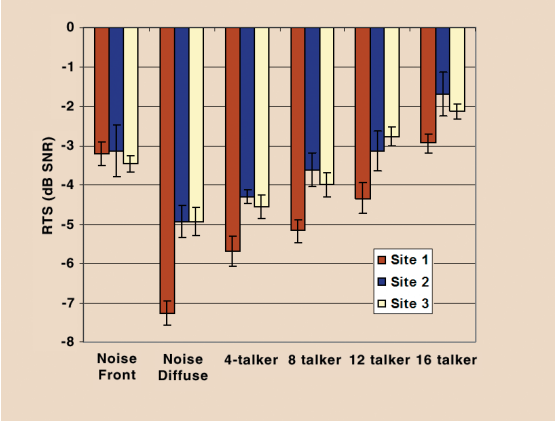

Figure 2: Average RTS (with standard error) from each site

CONCLUSIONS

Manipulations of basic presentation characteristics can have relatively large impacts on measured performance. Increasing the onset time of the masker improved performance by 2.5 dB. So mixing onset times when comparing aided and unaided conditions will inflate differences. The longer noise onset compensates for the artificial on/off presentation of the HINT noise and allows signal processing algorithms time to stabilize (as would be found in constant-noise environements).

Multi-source presentations of random noise are susceptable to interaural correlation effects, as shown by the difference in performance between correlated and uncorrelated noise sources. Perceptually, the correlation condition is similar to a diotic headphone condition with perceived location of the masker in the center of the head. The uncorrelated condition better approximates multiple external maskers and produces RTS almost 1 dB higher.

RTS with MTS varies in difficulty as would be expected if subjects can take advantage of the modulation in the maskers to extract target sentences in the quiet portions. The more talkers, the fewer quiet portions. Only one site is different from the others, and differences in soundfield volume are being investigated to explain this difference.

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 threhsolds in quiet and noise. J. Acoust. Soc. Am., 95(2), 1085-1099.
2. International Organization for Standardization (1992). Acoustics - Audiometric test methods - Part 2: Sound field audiometry with pure tone and narrow-band test signals. [ISO 8253-2]. Geneva: International Organization for Standardization.
3. Cox, R.M. & Moore, J.N. (1988). Composite speech spectrum for hearing aid gain prescriptions. Journal of speech and Hearing Reasearch, 31, 102-107.