Speech Recognition with High-Frequency Directed Audio

Ritvik P. Mehta, MD[1]; Sara L. Mattson, AuD[1]; Brian A. Kappus, PhD[2]
[1]California Hearing & Balance Center, La Jolla CA, [2]Turtle Beach Corporation, San Diego CA

ABSTRACT

Objectives: 1) To assess whether a second-generation HyperSound® Audio System (HSS), using ultrasound technology, improves speech recognition compared to a conventional acoustic speaker in subjects with mild to severe hearing loss. 2) To assess whether tuning the HyperSound system (HSS-tuned) to each subject's hearing loss results in further gain in speech recognition

Study Design: Single-blind, randomized cross over study.

Setting: Tertiary referral center.

Patients: Nine adult patients with mild to severe hearing loss with pure-tone average (PTA) of > 30 dB and word discrimination scores of < 80% in both ears.

Intervention(s): Subjects completed the AzBio Sentences test and the Consonant-Nucleus-Consonant (CNC) word test for each speaker type as well the tuned HyperSound system (HSS-tuned). The AzBio and CNC word tests were conducted in quiet at 70 dB. The AzBio was also conducted in noise at +10 SNR.

Main Outcome Measure: Speech recognition, as measured by the AzBio Sentences test and the CNC word test.

Results: Significant gains in speech understanding were observed using the HSS-tuned speaker versus both the conventional speaker and HSS without tuning for all test conditions at 70 dB. The mean AZBio scores for the HSS-tuned vs conventional speakers showed an average improvement in scores of 20% (p=0.015) in quiet, and 8.3% in noise (p=0.046). The CNC whole word test scores improved by an average of 23.8% (p=0.017) and CNC phoneme test scores improved by an average of 20.6% (p=0.021).

Conclusions: The second generation HSS system demonstrates highly significant improvement in speech recognition over conventional speakers in those with mild to severe hearing loss. Tuning the HSS system to the subject's hearing loss (HSS-tuned) demonstrates significant additional benefit in speech recognition across all test conditions.

Ritvik Mehta
California Hearing & Balance Center
Email: drmehta@sdearcenter.com
Phone: 858-909-0770
Website: www.calhns.com

INTRODUCTION

In 1963, Peter Westervelt was the first to theorize that highly directional receivers and transmitters may be constructed utilizing the nonlinearity of a medium to transmit an acoustic signal on a narrow beam over great distances.[1-3] This effect was initially observed experimentally in water in 1972[4] and, finally, in air by Bennett in 1974.[5] The effort to transform this research into a consumer product required both an increase in sound intensity and a reduction in distortion. Piezoelectric crystals[6] and films[7-9] were shown to produce airborne ultrasound with sufficient efficiency and intensity to be useful for parametric audio. Recently, the reduction of audible distortion through the application of high-speed digital signal processing and proprietary filtering techniques led to the commercialization of a parametric audio device.

Utilizing these innovations in a commercial product, HyperSound® Audio System (HSS) employed Polyvinylidene Fluoride (PVDF) emitters (speakers) (Figure 1) and proprietary digital signal processing to produce a focused column of high-fidelity sound. Considered a "directed audio" device, the HSS works by first electronically converting audible information onto ultrasonic frequencies, well-beyond the range of human hearing. The acoustic signal is reproduced using an emitter and transmitted in a beam of silent ultrasonic energy. The nonlinearity of air demodulates this acoustic signal, thus reproducing the audible information in a narrow beam, such that it is heard only by those in the targeted area. Unlike a conventional audio speaker, sound is not created omni-directionally at the speaker (emitter) surface but is created within a narrow, directional air column (Figure 2). Because sound can be transmitted in this narrow beam over a relatively long distance without a reduction in intensity, the HSS audio is not prone to degradations from ambient noise or reverberation. As a result, unlike conventional audio speakers, HSS maintains a high-fidelity audio signal at the same intensity level over a relatively long distance for listeners who are positioned in the beam.

In a recently published single blinded, randomized cross over study, HSS demonstrated highly significant improvement in unaided speech recognition over conventional speakers at 70 dB SPL, including in background noise, in those with mild to severe hearing loss.[10]

The present study was designed to assess whether tuning a second generation tunable HSS system to a subject's hearing loss by programming frequency specific gain results in further gain in speech recognition.

Figure 1. HyperSound® Audio System with PVDF Emitters

Figure 2. HSS Directional Air Column with Subject in the beam

METHODS

Participants
Participants were eligible for the study if they had a 4-frequency pure-tone average (PTA) of > 30 dB, and a maximum word recognition score of ≤ 80% in both ears. Nine adult patients who had mild to severe hearing loss were deemed eligible for the study.

Outcome Measures
The primary outcome was speech understanding ability, as measured by the AzBio sentences test and the Consonant-Nucleus-Consonant (CNC) word tests. Outcomes were measured in three test conditions: Second Generation HyperSound Audio System (HSS), Conventional Speakers (Phonic Ear audio system with two AT0578 speakers), and HyperSound Tuned (HSS-tuned) in which frequency specific gain was programmed using a fitting algorithm based on the subject's hearing loss. The order of each test condition was randomized and participants were blinded to the test condition. Each test condition was adjusted to achieve specified (70dB) output using a continuous 1kHz tone at the subject position using a handheld SPL meter (B&K 732A).

Table 1. φ All test scores are given as percent correct (referred to as clarity scores)

Test Condition	Conventional Speaker Scores (Mean±SD)	HSS Scores (Mean±SD)	P-value vs conventional speaker*	HSS-tuned Scores (Mean±SD)	P-value vs conventional speaker*
Az Bioφ					
70 dB, Quiet	38.4±33.2	46.7±41.5	0.036	58.4±34.5	**0.015**
70 dB, Noise	22.6±23.9	23.4±24.2	1	30.9±32.2	**0.046**
CNC Word Testφ					
70 dB (Whole Word)	22.4±23.9	30.2±28.2	0.028	46.2±32.7	**0.017**
70 dB (Phonemes)	38.7±29.0	44.7±33.4	0.11	59.3±33.6	**0.021**

*** p-values were obtained from the Wilcoxon Signed Rank Test**

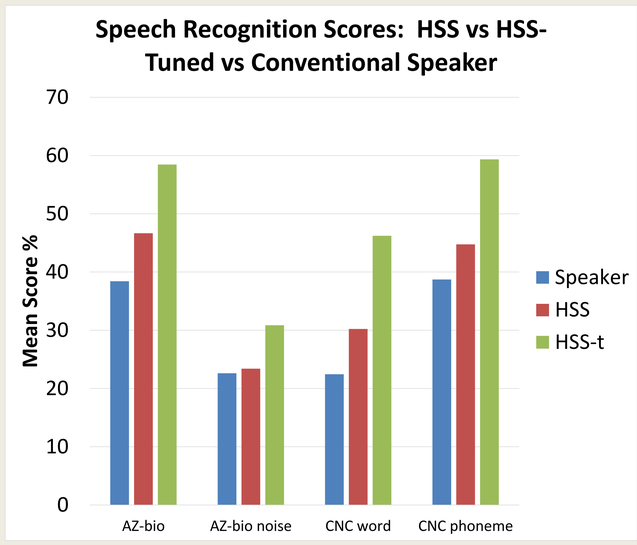

Chart 1. Clarity scores for each of the three test conditions. The benefit of tuning HSS to the subject's hearing loss is clearly demonstrated.

RESULTS

Statistically significant gains in speech understanding were observed using the HSS-tuned speaker versus both the conventional speaker and HSS without tuning for all test conditions at 70 dB. The mean AZBio scores for the HSS-tuned vs conventional speakers showed an average improvement in scores of 20% (p=0.015) in quiet, and 8.3% in noise (p=0.046). The CNC whole word test scores improved by an average of 23.8% (p=0.017) and CNC phoneme test scores improved by an average of 20.6% (p=0.021).

DISCUSSION

One reason that participants experienced greater speech intelligibility with the HSS may be due to the precise targeting of sound within a narrow beam. Unlike a conventional audio speaker, which disperses sound omni-directionally from the speaker surface, the HSS creates sound along and within a tight, directional air column. The precision targeting of the HSS significantly minimizes the effects of ambient noise and reverberation, so the sound beam maintains a clear, high-fidelity audible signal over a relatively long distance. Another explanation relates to the frequency response of the HSS compared to conventional speakers. It is possible that the HSS more effectively transmits a broader bandwidth signal with more high frequency information relative to the conventional speaker. Given the positive effects of additional high frequency energy on speech intelligibility,[11] it is possible the HSS is able to provide greater audibility of high frequency sounds, which contribute to improved speech understanding. Future studies are planned in an anechoic environment to further elucidate the advantages of directional audio over conventional speakers.

CONCLUSIONS

The second generation HSS system demonstrates highly significant improvement in speech recognition over conventional speakers in those with mild to severe hearing loss. Tuning the HSS system to the subject's hearing loss (HSS-tuned) demonstrates significant additional benefit in speech recognition across all test conditions.

REFERENCES

1. Westervelt PJ. "Parametric Acoustic Array". *J. Acoust Soc America 35(4): 535-537, April 1963*
2. Berktay HO and Leahy DJ. "Farfield performance of parametric transmitters". *J Acoust Soc America 55(3): 539-46, March 1974*
3. JLS and Beyer RT, "Experimental Investigation of an End-Fire Array," J Acoust Soc Am. 24, 1051-1054 (1962)
4. Willette and Muir. "Parametric Acoustic Transmitting Arrays". *J Acoust Soc Am* 1972
5. Bennett MB and Blackstock DT. "Parametric Array in Air" J Acoust Soc America 57(2): (562-568), March 1974
6. Gallego-Juarez JA, Rodriguez-Corral G, Gaete-Garreton L, "An Ultrasonic Transducer for high Power Application in Gasses" Ultrasonics Nov 1978
7. Sussner H, Michas D, Assfalg A, HunkLinger S, and Dransfeld K. "Piezoelectric Effect in Polyvinylidene Fluoride at High Frequencies" Physics Letters 45A(6) Nov 1973
8. Ambrosy A and Holdik K. "Piezoelectric PVDF films as ultrasonic transducers."00 *J Phys E Sci Instrum17: 856-859, 1984.*
9. Wang H and Toda M. "Curved PVDF Airborne Transducer" IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control 46(6) Nov 1999
10. Mehta RP, Mattson SL, Seitzman RL and Kappus BA. Speech Recognition in the Sound Field: Directed Audio vs Conventional Speakers. *Audiology Online*. August 31, 2015
11. Hornsby, BWY *(2004)*. The Speech Intelligibility Index: What is it and what's it good for? *Hearing Journal.* 57, 10, 10-17.