

The following technical paper has been accepted for presentation at the 2005 annual conference of the Rehabilitation Engineering and Assistive Technology Society of North America.

RESNA is an interdisciplinary association of people with a common interest in technology and disability. The organization's purpose is to improve the potential of people with disabilities to achieve their goals through the use of technology. RESNA promotes research, development, education, advocacy and provision of technology; and supports people engaged in these activities.

Functional Communication for Soft or Inaudible Voices: A New Paradigm

James Rothwell, DMA *, Dennis Fuller, PhD **

* Voicewave Technology Inc., Bridgeton, MO 63044

** Saint Louis University, Saint Louis, MO 63110

ABSTRACT

This paper introduces a proprietary electronic technology for improving the communication of soft or inaudible voices, contrasts it with prior technologies, and demonstrates its real-world functionality in extracting a fully intelligible whisper from an 85dB SPL noise-environment.

KEYWORDS

Soft voice; inaudible voice; speech; microphone; amplifier; aphonia; speech recognition; augmentative communication;

BACKGROUND

For more than fifty years electronic amplification has been tried as a means to boost soft or inaudible voices (SIV). Its unsatisfactory performance is due to the inability of microphones to pick up these voices and reject unwanted background noise: the softer the voice, the greater the difficulty.

Noise is always in our lives; so much so, we're usually unaware of it-- until it affects communication. Hearing our voice (auditory feedback) is a crucial component of the verbal output product: talkers monitor their output and adjust it accordingly. These adjustments may be subtle changes in pulmonary support, laryngeal tension, articulatory rate, etc., that the listener perceives as changes in voice projection, intelligibility of the message, or general vocal quality. The act of speaking is both a verbal and an auditory task in which persons with SIV find themselves at a disadvantage. Improved auditory monitoring could contribute to better production, but microphones aren't effective with SIV. Further, few microphone manufacturers provide specifications from which the outcome can be predicted.

There are many causes of SIV, including developmental disability, vocal abuse, and traumas such as high-level Spinal Cord Injury (SCI). SCI, for example, may impair the vocal mechanism or respiratory system, leaving a person unable to communicate in a work place or social setting, or direct his or her own medical care. Ordinary activities, like talking on the telephone, can become daunting challenges. If a person's SIV cannot somehow be separated from background noise, other empowering technologies such as computer speech recognition or voice-driven environmental controls are inaccessible to them.

"Inaudible" usually means "sound imperceptible by the ear". Noise isn't part of the definition, nor is the listener's hearing; yet, both influence the outcome. For this paper, "soft or inaudible voice" means a voice at less than normal output level -- without regard for its specific sound pressure level (SPL) -- that cannot be understood due to its softness, the listener's hearing ability, or because it is obscured by noise. All

three factors must be considered; a truly “functional” communication paradigm must include as a minimum, a talker, a listener, and the transmission medium or acoustic environment in which the communication occurs.

DEVELOPMENT

Improving the outcome for SIVs begins with prior technologies and their inadequacies. The basic principles of microphone design have been known for more than fifty years.(1) Unfortunately, little difference exists between the soft-voice and noise-rejection capabilities of today’s microphones and those of a half-century ago. Microphones employ one of two operating principles: directional discrimination or distance discrimination. Microphones with directional discrimination are most common, the prevalent types being omni-directional and cardioid (heart-shaped). Microphones with distance discrimination are called noise-canceling microphones and marketed at a premium.

Omni-directional microphones offer no noise rejection, picking up sound equally from all directions. Though ineffective in noise, they’re inexpensive and frequently used in low-cost systems.

Cardioid microphones offer a slight improvement, favoring sounds arriving at the front over sounds from the sides and rear. This matters little unless the environment is non-reverberant. Reverberant (acoustically reflective) environments distribute noise rather equally around a microphone, making directionality unimportant. Cardioids provide typically 6dB of noise rejection, often less.

Noise-canceling (NC) microphones discriminate against sounds on the basis of distance, rejecting sound originating at a distance (far field) while accepting sound originating near the microphone (near field). Their principal drawback is that rejection is undesirably frequency dependent: they reject best at low frequencies (100Hz) unimportant to intelligibility, but reject least in the critical region of 800Hz to 4KHz that is vital to intelligibility. Consequently, their success depends on the type and severity of noise in which they are used. Specifications, while seldom given, reveal worst-case noise rejection is typically 4dB to 6dB in the critical region, increasing by 6dB per octave as frequency decreases -- a most severe limitation if there is competing noise in the critical region.

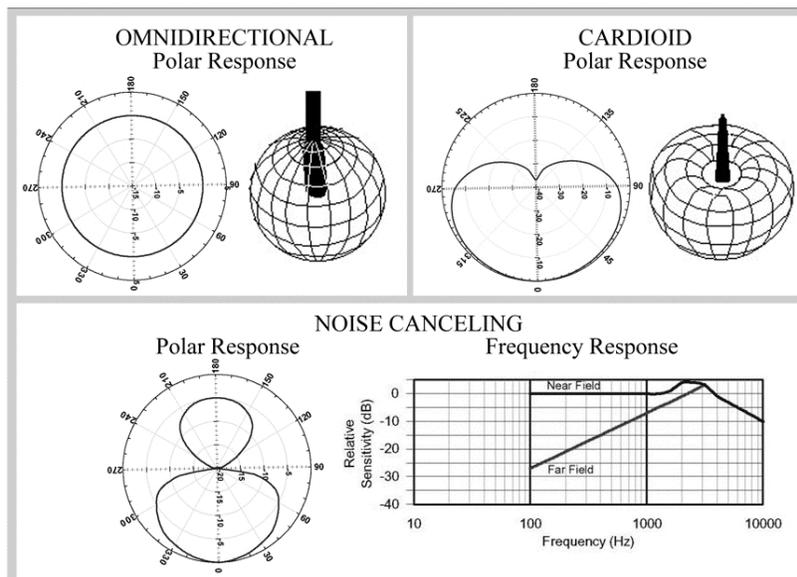


Figure 1. Microphone Patterns

To have functional, real-world communication, as used here, requires that a person able to converse in any situation encountered in daily life. Although quiet residential levels may range from 50dB to 60dB SPL, an active household can top 80dB SPL. Workplaces and offices seldom have noise levels under 65dB to 75dB SPL, and industrial, school, and workshop environments can exceed 95dB SPL. Most surprisingly for those who would direct their own medical care, one hospital noise study documented jackhammer-like noise levels of 113dB SPL inside patient rooms during the morning shift change.(2)

To understand the requirements for accommodating SIVs we examine the speech signals themselves. In normal adult speech, low frequency components have the highest amplitude, being the chief contributors to the average speaking level of 55dB to 65dB SPL. These components, though, contribute little to intelligibility. Conversely, the speech components in the critical 800Hz to 4KHz region are lowest in amplitude -- typically 20dB to 30dB less than low frequency components -- but are essential to intelligibility. Since a “normal” voice emits these critical components at 25dB to 35dB SPL, they are obviously far lower in SIVs, making them easily obscured as noise rises or voice declines. The talker’s diminished auditory feedback usually causes them to overdrive pulmonary, articulation and voicing characteristics to the point of hyper-function. Hyper-functional characteristics result in distorted verbal output, requiring significant energy consumption -- often to the point of premature speaker fatigue. This in turn leads to diminished intelligibility and/or loss of voicing capability. Auditory feedback is vital.

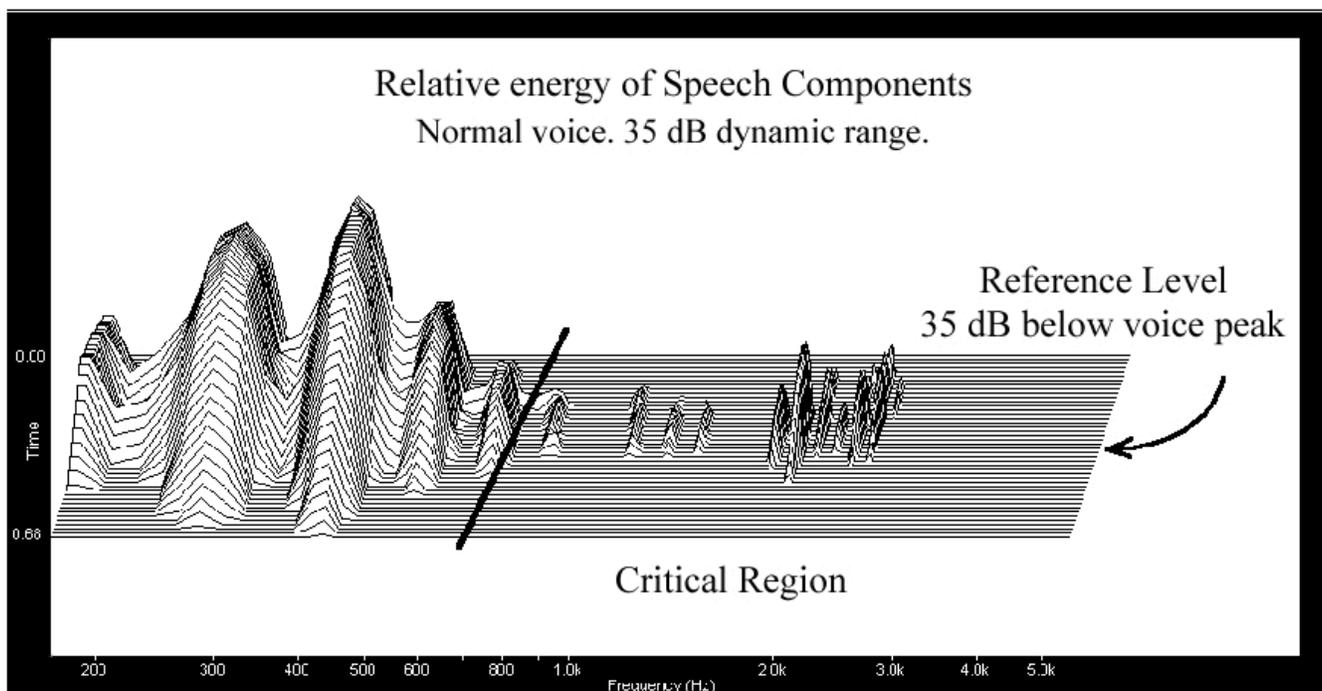


Figure 2. Speech Components

Traditional microphones are designed for positive signal-to-noise environments: more voice than noise is required for satisfactory performance. Real-world conditions, though, dictate that often more noise is present than voice. Functional SIV communication, then, requires accommodating voices far weaker than noise -- by perhaps as much as 70dB to 80dB -- while rejecting noise across the entire audio spectrum.

The Voicewave Technology Inc. (Voicewave) proprietary speech technology was designed to these requirements. It's a three-fold approach, comprising uniquely effective microphone architecture, an electronic signal processor, and an auditory feedback mechanism to the talker. Together, they provide over 99% broadband rejection of noise, aid speech production through improved self-monitoring, and have the sensitivity to pick up signals at or below the threshold of hearing. This dramatic increase in performance enables persons with SIVs to communicate virtually without regard for their noise environment, using motor articulation and pulmonary support levels that produce the most efficient performance levels. The resulting verbal output can extend communication time, while improving the verbal output quality and reducing talker fatigue. Further, the technology opens up other opportunities to enhance communication, such as speech clarification and speech recognition --subjects for future papers.

Figure 3 illustrates the superior performance of Voicewave technology over a standard cardioid microphone in picking up a whisper in a reverberant environment of 85dB SPL noise. While noise totally obscures the whisper with the cardioid, Voicewave's technology extracts a fully intelligible whisper.

SIGNAL SOURCE: Whisper in 85 dB Noise

A. Cardioid Microphone

B. Voicewave Proprietary Technology

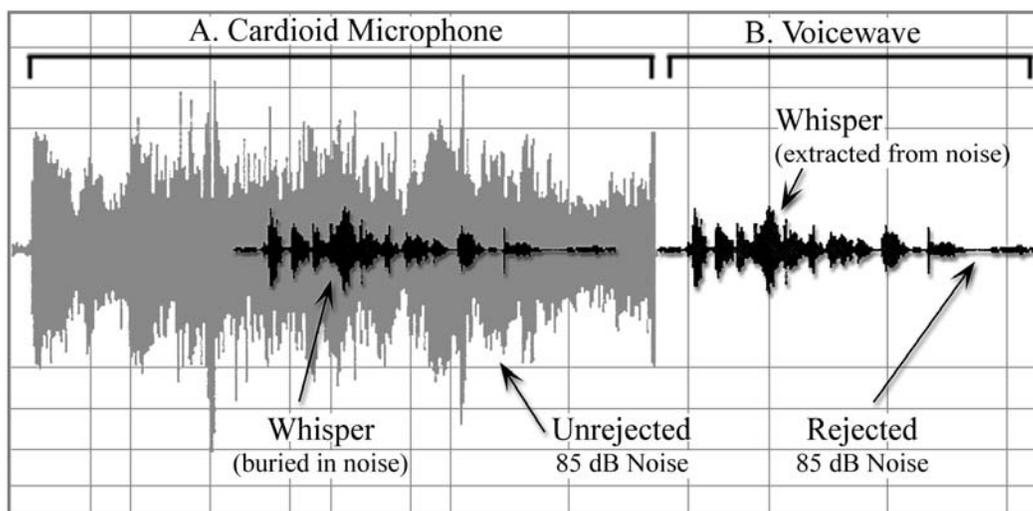


Figure 3. Noise Rejection

The benefits of this technology are significant for SIV. The nearly total rejection of background noise enables boosting SIVs to a far greater extent than allowed by other technologies. A side benefit is reduction of stress, since less effort is expended to produce speech, and improved self-monitoring -- essential for coordinating speech processes. The Voicewave processor incorporates spectrally enhanced auditory feedback that improves "hearability" in noise, increasing the auditory feedback's effectiveness by 10dB to 15dB without increasing its volume level.

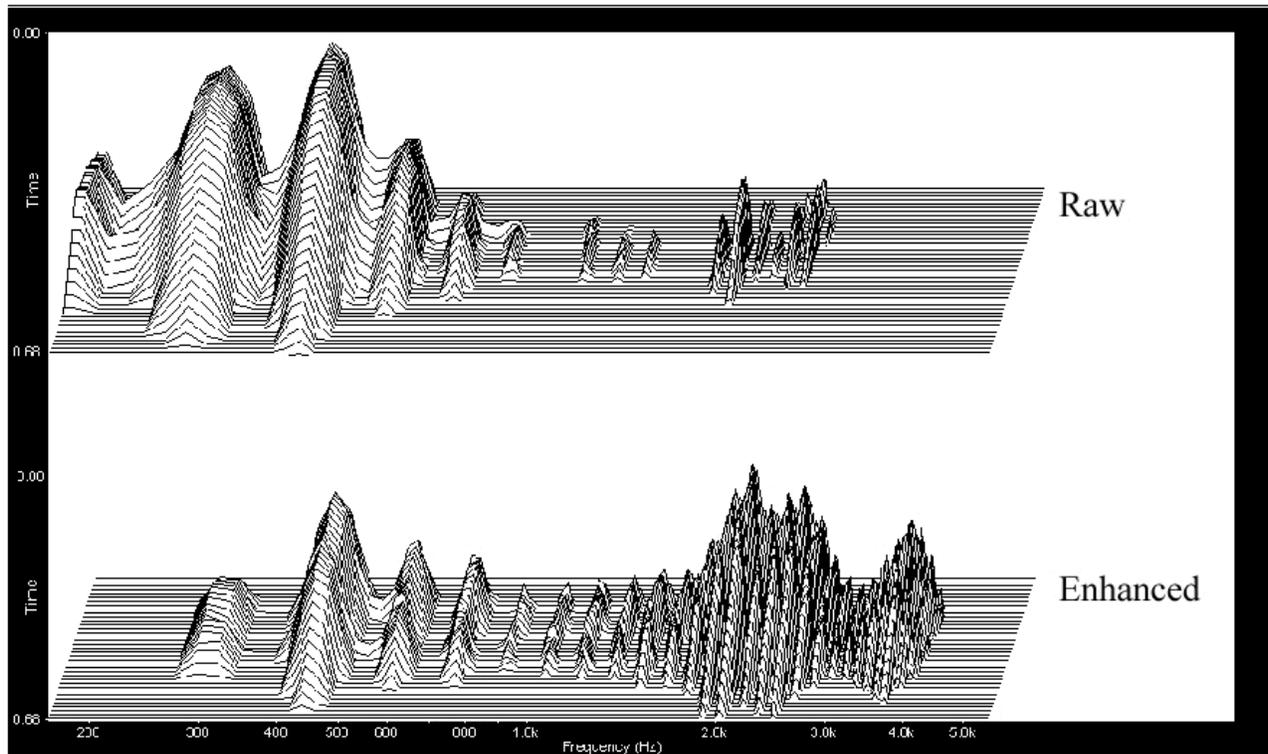


Figure 4. Raw vs. Enhanced Speech

CONCLUSION

The technology offers persons with SIV a new means of functional communication, without regard for their communications environment, plus the benefits of reduced effort and stress. It also meets the durability requirements of physically challenged users, and its microphone technology is water-resistant.

1. Olson, Harry F. (1957). *Acoustical Engineering*. Princeton, NJ: D. Van Nostrand Company, Inc.
2. Cmiel, C. A., Karr, D. M., Gasser, D. M., Oliphant, M. L., Neveau, A. J. (2004). Noise Control: A Nursing Team's Approach to Sleep Promotion. *American Journal of Nursing*, 104 No. 2, 40-48.