JANUARY 01 1955

# The Noise-Cancelling Headset—An Active Ear Defender

Elvin D. Simshauser; Mones E. Hawley



*J. Acoust. Soc. Am.* 27, 207 (1955)
https://doi.org/10.1121/1.1917899

**H9. The Noise-Cancelling Headset—An Active Ear Defender.** Elvin D. Simshauser* and Mones E. Hawley, *Special Devices Engineering, Radio Corporation of America, Camden, New Jersey.*—The purpose of this device is to reduce ambient noise at the ear by using a headset to generate a sound pressure equal in magnitude and opposite in phase to the noise. This paper reports the results of pure-tone experiments. A microphone was mounted on each earphone of a conventional military headset. A two-channel system was used in which the signal from each microphone was passed through an amplifier and a phase-shifting network and was then applied to its earphone. The subject wore the headset in a pure-tone field and adjusted the amplifications and phase shifts for minimum loudness. The signal produced by each earphone was then substantially equal in magnitude and opposite in phase to the sound coming through and under its earcap. The proper amplification and phase shift are frequency-dependent and vary among subjects. Eight complete measurements were made on two subjects. The mean values of the two parameters were computed for each frequency. Calculations indicate that a device having these mean characteristics would provide a noise reduction between 100 and 1200 cps that averages 10 db more than is provided by the earcap alone.

* Now on leave of absence with the U. S. Army.

## Session I. Speech and Hearing

John Webster, *Chairman*

### *Contributed Papers*

**I1. Electrical Synthesizer of Continuous Speech.** K. N. Stevens and R. P. Bastide, *Acoustics Laboratory, Massachusetts Institute of Technology, Cambridge, Massachusetts,* and C. P. Smith, *Air Force Cambridge Research Center, Cambridge, Massachusetts.*—An electrical speech synthesizer controlled by teletype equipment is described. The basic components of the synthesizer are four simple tuned circuits each of which represents one of the resonances of the human vocal tract. The resonant frequencies of three of the tuned circuits are set by relay circuits that are controlled by data coded on punched tape. The punched tape also controls the intensity and pitch of a buzz generator that excites the variable-tuned circuits, and modulates a noise generator that is filtered in various ways to generate fricative or stop consonants. The coded data are read from the tape at about the phonemic rate, with a rate of information transmission of about 35 bits per second. Smoothing circuits are incorporated in the synthesizer to provide smooth formant transitions between sounds. The generation of individual sounds, words, and sentences by the synthesizer is demonstrated.

**I2. Influences of Variations in Speech Intensity and Other Factors upon the Speech Spectrum.** J. C. R. Licklider, *Massachusetts Institute of Technology, Cambridge, Massachusetts,* and Mones E. Hawley and Robert A. Walkling,* *Radio Corporation of America, Camden, New Jersey.*—In calculating speech intelligibility by methods such as French and Steinberg's, and in other phases of the designing of communication systems, one uses the long-time-average power density spectrum of speech. Several determinations of this spectrum have been made previously, but in each case the talkers talked at a fixed level. The present observations are aimed primarily to determine how the shape of the spectrum differs from "normal" when, as is necessary in some situations, the talker raises his voice considerably above normal level. Secondary questions concern the amount of variation among talkers and among speech samples and the length of sample required for stable measurement. Results, obtained with two different spectrum analyzers, are presented. As the speech level is increased, the spectrum changes noticeably—enough to affect the outcome of intelligibility calculations. The middle- and high-frequency components are raised much more than the low-frequency components. The spectrum of normal speech has its maximum ordinate at 100 or 200 cps, whereas the spectrum of very intense speech has its maximum ordinate at 700 or 800 cps.

* Now at Harvard University, Cambridge, Massachusetts.

**I3. Automatic Extraction of Formant Frequencies from Continuous Speech.** James L. Flanagan, *Air Force Cambridge Research Center, Cambridge, Massachusetts, and Acoustics Laboratory, Massachusetts Institute of Technology, Cambridge, Massachusetts.*—An electronic apparatus for extracting formant frequencies from continuous speech is described. The short-time speech frequency spectrum or its second derivative with respect to frequency is sampled rapidly and periodically by scanning the outputs of a thirty-six channel filter set. The time function thus produced is analyzed to determine the spectral maxima or formants. The points in time at which maxima occur are separated and converted into voltages by a further sampling and smoothing operation. These voltage outputs correspond to the formant frequencies as functions of time. A vowel segmenting circuit is incorporated in the input to restrict analysis to the vocalic portions of the input speech if desired. Some preliminary results are presented.

**I4. An Analysis of Clipped Speech.** Friedrich Vilbig, *Air Force Cambridge Research Center, Cambridge 39, Massachusetts.*—The application of infinite clipping to speech has only a small influence on intelligibility. The reason for this is that clipping the original speech-spectrum only causes tolerable distortion. Since the mathematical analysis for many frequencies becomes involved to a high degree of difficulties, the theoretical investigation was restricted to 2 frequencies, $a\cos\alpha t+b\cos\beta t$ e.g., representing 2 harmonics of the pitch frequency. The clipping is considered to be a rectification of a rectangular curve $i=f(x)$ where $i=0$ for $-\pi<x<0$ and $i=A$ for $0<x<+\pi$. Only the range $-\pi\cdots0\cdots\pi$ is used. The resulting amplitudes are expressed as infinite sum of a product of 2 Bessel functions. The results of this calculation are confirmed by an experimental analysis. Highest distortions appear for $a=b$. Considering $\alpha=(\alpha+\beta/2)+(\alpha-\beta/2)$ and $\beta=(\alpha+\beta/2)-(\alpha-\beta/2)$, the 2 unclipped frequencies represent 2 sideband frequencies of a suppressed carrier frequency, $(\alpha+\beta/2)$. By infinite clipping, the frequencies $n(\alpha+\beta/2)\pm(\alpha-\beta/2)$ are the most important ones in amplitude, where $n=1, 3, 5\cdots$. Thus, speech clipping, as applied to vowels, causes distortions mostly of frequencies above 3000 cps. e.g. $n=3$ or higher. But if the clipped speech is put through a 3000-cps low-pass filter, the difference between original speech and clipped speech is not very important. Since $\alpha$ and $\beta$ and all distortion frequencies are harmonics of the pitch frequency, the good intelligibility of clipped speech is quite reasonable. An experimental analysis of several vowels and consonants has proven this theory.