AUGUST 03 2005

Human or Machine?—a Study of Voice Naturalness

L. G. Kersta; P. D. Bricker; E. E. David, Jr.

J Acoust Soc Am 32, 1502 (1960)
https://doi.org/10.1121/1.1935196

the effects of peak clipping on speech intelligibility have been investigated, the influence of other forms of nonlinear distortion are not as well known. In the present study, we introduced specified amounts of second-order distortion into PB word lists. A square law circuit generated the distortion products. The playback of the previously-recorded PB word lists was split into two channels which fed a dual-track tape recorder. One channel contained the distortion generator. Thus, while the original clean speech was simply dubbed onto one track, the other track recorded only the second-order distortion products of that same speech. On playing back the dual-track tape, any degree of distortion could be obtained by varying the relative amplitudes of the two channels. Intelligibility score (percent of words correct) was investigated as a function of four degrees of distortion, nominally termed 0, 30, 50, and 100% distortion. Findings showed that normals and conductive losses withstood the effects of distortion far better than those with sensori-neural impairments and that the intelligibility score of the sensori-neural was a decreasing monotonic function of the degree of neural impairment.

G9. Reduction of Signal-to-Noise Ratio Induced by Distortion in Speech Transmission Systems. EDITH L. R. CORLISS, MICHAEL KOBAL (Student Assistant, now at Boston University), AND SUSAN E. BURGHORN (Student Assistant, now at University of Chicago), *National Bureau of Standards, Washington, D. C.*—A study has been made of the influence of transient and intermodulation distortion on the transmission of speech occurring in the presence of a background noise having pronounced transient characteristics (cafeteria noise). This paper describes objective measurements of the power-time distribution in speech and noise, separately and in combination, for the sound coming from a sound source of high quality and for the sound coming from the earphones of hearing aids whose microphones were exposed to the sound source. The nature of the transmitted signal will be demonstrated. (This work was carried out under sponsorship of Veterans Administration.)

G10. An Approach to the Evaluation of Vocoded Speech. WILLARD F. MEEKER, *Radio Corporation of America, Camden, New Jersey,* AND CALDWELL P. SMITH AND WILLIAM R. SMITH, *Communication Sciences Laboratory, Air Force Research Division, L. G. Hanscom Field, Bedford, Massachusetts.*—Distortions peculiar to speech analysis-synthesis devices, such as the vocoder, introduce new problems in the determination of performance standards. Intelligibility can be determined in the laboratory, but standardized methods for the measurement of the characteristics of speech quality and talker recognition have not yet been established. The researcher or engineer has little to guide him in the inevitable compromises involved in the development and design of such devices and in the selection of optimum data rates. In an effort to provide such guidance, a survey was first made of certain potential military users. Speech samples obtained by filtering, telephone transmission, and vocoder processing were presented for judgments regarding intelligibility, speech quality, talker recognition, and adequacy for normal use. This paper presents the results of these surveys, along with data regarding intelligibility and talker recognition obtained in the laboratory. The speech samples used in the survey will serve as a basis for comparison in laboratory judgments to determine approximate boundaries for those parameters affecting the characteristics considered.

G11. Human or Machine?—a Study of Voice Naturalness. L. G. KERSTA, P. D. BRICKER, AND E. E. DAVID, JR., *Bell Telephone Laboratories, Inc., Murray Hill, New Jersey.*—Critical examination of human voiced speech indicates that both sustained vowels and connected speech often show considerable variability in the spacing of successive glottal pulses. Such fluctuations are too rapid to be reproduced by a low-frequency parametric pitch signal as classically used in vocoders. What is the perceptual significance of these fluctuations? A preliminary investigation of this question utilized sustained vowels spoken by two talkers. The fluctuations in these utterances were removed by digital computer processing. These samples were used, together with the originals, to form A-B pairs, each pair containing one "machine" and one "human" utterance. Subjects were presented with the A-B pairs and asked to identify the one made by the human. Naive subjects were able to perform almost perfectly in this task. Since rapid period-to-period fluctuations were the distinguishing characteristic of the human samples, this result is evidence for the role of such fluctuations in voice naturalness.

G12. Speech Detection in Noise. PAUL S. VENEKLASEN AND JERRY P. CHRISTOFF, *Western Electro-Acoustic Laboratory, Inc., Los Angeles 49, California.*—Exploratory work to optimize the signal-to-noise ratio and the wearability of speech detection devices for use in high-altitude flight has led to quantitative determination of the relative capabilities of head contact devices in relation to acoustic detectors. One of the discoveries of the exploratory work was that an accelerometer fastened to an upper front tooth will detect highly intelligible speech and discriminate strongly against external noise. Data is presented on the acceleration spectrum of the teeth as a result of speech and external noise. The signal-to-noise ratio of the tooth microphone is compared to other speech reception systems under the same conditions. Refinement of the attachment of the accelerometer to the teeth is discussed. (This research was supported by U. S. Air Force contract.)

Session H. Panel on Instrumentation Problems in Missile Noise

T. J. SCHULTZ, *Chairman and Panel Moderator*

Invited Papers

H1. Review of the Noise Generation of Rockets and Jets. KEN ELDRED, *Western Electro-Acoustic Laboratory, Inc., Los Angeles 49, California.*—The recent emphasis on the study of jet-noise generation has resulted in a large body of data giving the acoustical power, spectra, and directivity of the noise radiated from jets which vary in nozzle diameter from ½ to 40 in., in velocity from 500 to 6500 ft/sec, and in total temperature from 525 to 4500°R. This paper examines the relationships between some of these acoustical data and the various accompanying jet parameters in an effort to define similarity laws which apply to all jets studied. These