The Sound Spectrograph

W. Koenig, H. K. Dunn, and L. Y. Lacy

Citation: The Journal of the Acoustical Society of America **18**, 244 (1946);
View online: https://doi.org/10.1121/1.1902419
View Table of Contents: http://asa.scitation.org/toc/jas/18/1
Published by the Acoustical Society of America

Program of the Thirty-First Meeting of the Acoustical Society of America

Hotel Pennsylvania, New York, New York

May 10–11, 1946

Introduction (5 min.) - R. K. Potter.

1. The Sound Spectrograph.—W. Koenig, H. K. Dunn, and L. Y. Lacy, *Bell Telephone Laboratories, Inc., New York, New York* (10 min.).—The sound spectrograph is a wave analyzer which produces a permanent visual record showing the distribution of energy in both frequency and time. This paper describes the operation of this device and shows the mechanical arrangements and the electrical circuits in a particular model. Some of the problems encountered in this type of analysis are discussed, particularly those arising from the necessity for handling and portraying a wide range of component levels in a complex wave such as speech. Spectrograms are shown for a wide variety of sounds, including voice sounds, animal and bird sounds, music, frequency modulations, and miscellaneous familiar sounds.

2. Translators with External Phosphors. Homer Dudley and Otto O. Gruenz, Jr., *Bell Telephone Laboratories, Inc., New York, New York* (10 min.).—Two translators are described which display instantaneously on a moving external phosphor the essential characteristics of sound waves, such as speech. There is thus presented to the eye a parade of speech patterns as though a printed line were passing by. The pattern of intensity *vs.* frequency *vs.* time remains in view long enough to permit the eye to observe pattern groups as an aid to speech interpretation. Typical applications are portraying the phonetics of speech and aiding the deaf to understand speech and to build up their own speech. The goal in the speech pattern portrayal is to show the stronger speech components simply but accurately as a function of frequency and time. The speech is equalized to bring out the stronger components over the frequency range. Next automatic volume control is applied to iron out level differences on a time basis. Then the 3500 cycle speech band is analyzed by twelve filters of about 300 cycles each and the resultant component powers applied to the excitation of a phosphor. The smaller translator employs twelve incandescent grain-of-wheat lamps for phosphor excitation, generating a pattern 1″ high and 5″ long and the larger translator sets up a pattern 7″ high and 47″ long. The patterns remain in the field of view for about two seconds.

3. The Cathode-Ray Translator. R. R. Riesz and L. Schott, *Bell Telephone Laboratories, Inc., New York, New York* (10 min.).—A system has been developed whereby speech analysis patterns are made continuously visible on the moving luminescent screen of a special cathode-ray tube. The screen is a cylindrical band that rotates with the tube about a vertical axis. The electron beam always excites the screen in the same vertical plane. Because of the persistence of the screen phosphor and the rotation of the tube, the impressed patterns are spread out along a horizontal time axis so that speech over an interval of a second or more is always visible. The upper portion of the screen portrays a spectrum analysis and the lower portion a pitch analysis of the speech sounds. The frequency band up to 3500 cycles is divided into 12 contiguous sub-bands by filters. The average speech energy in the sub-bands is scanned and made to control the excitation of the screen by the electron beam which is swept synchronously across the screen in the vertical direction. A pitch detector produces a d.c. voltage proportional to the instantaneous fundamental frequency of the speech and this controls the width of a band of luminescence that the electron beam produces in the lower part of the screen. The translator has been used in a training program to study the readability of visible speech patterns.

4. The Portrayal of Visible Speech. J. C. Steinberg and N. R. French, *Bell Telephone Laboratories, Inc., New York, New York* (10 min.).—This paper is concerned with the portrayal of speech in the form of visible patterns that one might learn to read as readily as one learns to understand spoken speech. The portrayal is to be accomplished automatically from the sound waves without a serious time lag between spoken word and visible patterns. In principle, the method attempts to perform a frequency-intensity-time analysis of the sound somewhat analogous to that performed by the ear, and to present the results in an orderly manner to the eye. The problem of obtaining legible patterns is one of choosing suitable frequency and time intervals in which to measure the intensity and suitable frequency, intensity, and time scales for the visual portrayal. The choice involves the characteristics of both the eye and the voice. The paper discusses objectives and requirements based on these characteristics, describes the patterns in current use and explores the possibilities and implications of certain modifications that might be made in the present portrayal.

5. Basic Phonetic Principles of Visible Speech. G. A. Kopp and H. C. Green, *Bell Telephone Laboratories, Inc., New York, New York* (10 min.).—The phonetic principles basic to the legibility and interpretation of visible speech are discussed in this paper along with the unique characteristics which differentiate the patterns of individual sounds and sound groups. The visible patterns are interpreted in terms of physiological phonetics and the types of modulation used by the speaker in producing the sounds. A threefold comparative classification is developed, including