

Differential Sensitivity to Synthetic Speech Sounds within and between Phoneme Categories

K. S. Harris, A. M. Liberman, H. Hoffman, et al.

Citation: [The Journal of the Acoustical Society of America](#) **28**, 760 (1956); doi: 10.1121/1.1905011

View online: <https://doi.org/10.1121/1.1905011>

View Table of Contents: <https://asa.scitation.org/toc/jas/28/4>

Published by the [Acoustical Society of America](#)

JASA
THE JOURNAL OF THE
ACOUSTICAL SOCIETY OF AMERICA

Special Issue:
Additive Manufacturing and Acoustics

Read Now!

developed by the Centre National d'Etudes des Télécommunications de Paris for evaluating telephone sets objectively. This objective evaluation procedure replaces the subjective evaluations hitherto in normal use. The system consists of an artificial voice that is mechanical or electronic depending on the manner of use (tests in factories or measurements at the home of the subscriber), an artificial mouth, and an artificial ear (a complete and simplified model). The characteristics of this equipment duplicate those of the corresponding human organs as determined for a large number of subjects. The artificial mouth is set in an artificial head. The average spectrum of the artificial voice reproduces that of the natural voice. The artificial ear contains two elements: One is geometrical and mechano-acoustical, and simulates the acoustical impedance of the ear in such a manner that the sound pressure produced by a receiver applied against its cavity is the same as the sound pressure produced in the case of a similar application against an average external auditory meatus. The second, purely electrical, element gives to the detector system of the artificial ear a transfer function that is analogous to that of a natural ear. Experience has shown a very satisfactory correspondence between results of measurements performed with the described equipment and those obtained by a team of well-trained human subjects.

DC3. Description d'ensemble des principes physiques et des appareils techniques expérimentés par le Centre National d'Etudes des Télécommunications, pour le contrôle objectif des appareils téléphoniques et sa substitution au contrôle subjectif jusqu'ici en service normal. Le système utilisé comporte une voix artificielle, mécanique ou électronique (suivant le mode d'emploi: vérification en usine, ou mesure au domicile de l'abonné), une bouche artificielle et une oreille artificielle (modèle complet et modèle simplifié). Les caractéristiques de ces appareils reproduisent celles des organes correspondant qui ont été déterminées sur un grand nombre de sujets. La bouche artificielle est incorporée à une tête artificielle. Le spectre moyen de la voix artificielle reproduit celui de voix naturelle. En ce qui concerne l'oreille artificielle, elle comporte deux éléments, l'un géométrique et mécano-acoustique qui simule, indirectement, l'impédance acoustique de l'oreille, de telle façon que la pression acoustique, produite par un récepteur appliqué contre son pavillon, soit la même que la pression produite dans le cas d'une application semblable contre le pavillon d'une oreille naturelle moyenne. Le second élément, purement électrique de l'oreille, donne au système détecteur de l'oreille artificielle une loi de fonctionnement analogue à celle de l'oreille naturelle. L'expérience a montré une correspondance très satisfaisante entre les résultats des mesures exécutées avec le dispositif décrit et avec une équipe d'opérateurs bien entraînés.

DC4. Differential Sensitivity to Synthetic Speech Sounds within and between Phoneme Categories. K. S. HARRIS, A. M. LIBERMAN, H. HOFFMAN, AND B. C. GRIFFITH, *Haskins Laboratories, New York, New York, and University of Connecticut, Storrs, Connecticut.*—The use of synthesizers provides an opportunity to vary speech-like sounds in small steps along a single acoustic continuum. We find that at some points on such a continuum large changes in the acoustic stimulus have no effect on phoneme identification, while at other points small changes cause the listener's identification to shift abruptly from one phoneme to another. Casual observation led us to suspect that there might be related discontinuities in the discriminability of these sounds—that is, that discrimination would be less sharp, other things equal, between sounds in the same phoneme category than between sounds which lie on opposite sides of a phoneme boundary. This effect might be related to the difficulties that linguists often experience in hearing certain differences among the sounds of an exotic language. Discrimination and identification functions were obtained for a series of stimuli which differed

in the second-formant transition. In one part of the experiment these stimuli were presented singly to listeners for identification as *b*, *d*, or *g*; in another part, the discriminability of the stimuli was measured by an *ABX* technique. It was found that discrimination was, indeed, better in the vicinity of phoneme boundaries than it was near the middle of a category. The obtained discrimination function was quite close to a function predictable from the identification judgments on the extreme assumption that the listeners were able to discriminate the sounds only to the extent that they could differentially identify them as *b*, *d*, and *g*. [This work was supported in part by the Carnegie Corporation of New York, and in part by the Department of Defense in connection with Contract DA49-170-SC-1642.]

DC5. The Modern Russian Sentence Intonation as a Linguistic Structure. J. ERNA JURGENS BUNING AND CORNELIUS H. VAN SCHOONEVELD, *University of Leiden, Leiden, Netherlands.*—A code of intonational features was found which, if substituted for one another in the message, entail a change in the information carried by the sentence, *viz.*, (a) a fall of a fifth to the tonic of the underlying musical key, *i.e.*, a decrease of fundamental frequency in the proportion 3:2, from the immediately preceding syllable to the last stressed syllable *characterizes the statement*; (b) a rise of a third from the underlying tonic, *i.e.*, an increase of frequency in the proportion 5:6 from the last stressed syllable to the next syllable, *characterizes the question* (when no interrogative pronoun or particle is used). In a large sentence additional constant characteristic intervals (*i.e.*, proportions in frequency) precede. The interval constants characterizing larger statements are contextual variants of the basic type: the fall of a fifth to the tonic. Sentences containing contrastive stress have larger intervals; their form can easily be derived from the intonation patterns of comparable sentences without contrastive stress. The tone on which a single syllable is spoken can be interpreted only by juxtaposition with the other syllable tones of the sentence. Thus the musical concepts of key, degree, and interval, which superimpose upon the chromatic spectrum of musical tones a limit of choice, become relevant for the interpretation of sentence intonation.

DC6. Realization of Prosodic Features in Whispered Speech. WERNER MEYER-EPLER, *Institut für Phonetik und Kommunikationsforschung der Universität, Bonn, Germany.*—Several investigators have dealt with the problem of whispering in tone languages. As the perception of pitch changes demands a certain degree of periodicity (voice) within the speech sounds, it was difficult to see how tonal features could be preserved in whispered speech. Observations with Chinese speaking people, however, had indicated that there is no remarkable loss of intelligibility if Chinese is whispered instead of being spoken normally. The same result was found in Indo-European languages where no tones are used as phonemically distinctive elements. Intonation here belongs to the suprasegmental (prosodic) features, different types of intonation being correlated with different speech situations. Experiments were carried out by using a visible-speech analyzer which demonstrate that changes of pitch in normal (voiced) speech are replaced by shifts of some formant regions in whispered speech, sometimes accompanied by considerable broadening of the formant regions due to an increased velocity of the air stream passing the glottal fissure.

DC7. Preliminary Investigation of the Acoustical Characteristics of Inter-Phonemic Transitions. MARTIN C. SCHULTZ, *Speech and Hearing Center, Hospital of the University of Pennsylvania, Philadelphia, Pennsylvania.*—The study employed selected vowels, each combined with prevocalic or post-vocalic consonants in consonant-vowel or vowel-consonant syllables. Each combination was recorded non-