TU/e EINDHOVEN UNIVERSITY OF TECHNOLOGY

Contributions of voice-source and vocal-tract characteristics to speak identity

Citation for published version (APA):
Eggen, J. H., Nooteboom, S. G., & Houtsma, A. J. M. (1992). Contributions of voice-source and vocal-tract characteristics to speak identity. *Journal of the Acoustical Society of America*, *92*(4).

Document status and date:
Published: 01/01/1992

Document Version:
Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.tue.nl/taverne

Take down policy
If you believe that this document breaches copyright please contact us at:

openaccess@tue.nl

providing details and we will investigate your claim.

Download date: 04. Oct. 2023

(Walter Reed Army Med. Ctr., Army Audiol. and Speech Ctr., Washington, DC 20307-5001)

In a recent article [K. W. Grant and L. D. Braida, J. Acoust. Soc. Am. **89**, 2952–2960 (1991)] it was demonstrated that spectrally different bands of speech with equal articulation index (AI) scores provided approximately equal auditory-visual sentence recognition when combined with speechreading. Given that different parts of the frequency spectrum provide different segmental cues for consonant and vowel recognition, current models of auditory-visual integration [e.g., L. D. Braida, J. Exp. Psychol. **43A**, 647–677 (1991)] would predict that some spectral regions of speech are more complementary to speechreading than others. This raises the possibility that the findings of Grant and Braida are attributable, at least in part, to suprasegmental cues that were transmitted equally well by the different spectral bands tested. To test this possibility, the identification of syllable number, syllable stress, sentence intonation, and phrase boundary location was assessed under six approximately equal AI filter conditions similar to those evaluated by Grant and Braida. The results indicate that syllable number and syllable stress are perceived best through high-frequency bands, intonation is perceived best through low-frequency bands, and phrase boundary location is perceived equally well throughout the speech spectrum. These results are discussed in terms of the importance of different spectral regions of speech for the recognition of suprasegmental cues, and how this may relate to overall speech intelligibility. [Work supported by NIH Grant No. DC00792.]

11:15

1aSP13. Estimation versus word recognition of monosyllabic words. Rory A. DePaolis (Dept. of Special Education, Southeastern Louisiana Univ., Box 879, SLU Station, Hammond, LA 70402) and Tom Frank (Penn State Univ., University Park, PA 16802)

The relation between word recognition and the estimate of the percentage of words understood was examined for monosyllabic words. Six-hundred and sixteen words were presented to 30 subjects in four signal-to-noise ratios (SNR) under five high pass, five low pass, and in the unfiltered condition. In each condition the subjects listened to 14 words with the carrier phrase "you will write" preceding each word. Subjects were instructed to write down each word, and at the end of each set of 14 words, to estimate the percentage of words that they understood in 7.5% increments. The subjects were not able to reliably estimate the words that they correctly understood based upon their identification of the words. Confusion matrices for 10 of the 30 subjects show the filter conditions and phonemic classifications that subjects consistently misjudged as having correctly understood.

11:30

1aSP14. Contributions of voice-source and vocal-tract characteristics to speaker identity. J. H. Eggen, S. G. Nooteboom, and A. J. M. Houtsma (Inst. for Perception Res./IPO, P. O. Box 513, NL 5600 MB Eindhoven, The Netherlands)

A study was done on whether a voice-source model, in particular the Liljencrants–Fant model, codes information that is used by listeners to identify speakers by their voices. Automatic analysis/resynthesis techniques were used to generate so-called hybrid vowels for which the voice-source characteristics of one speaker are combined with the vocal-tract characteristics of another speaker. Listeners who were trained to recognize the speakers by their natural utterances, had to indicate which speaker they thought had produced the hybrid stimulus. It was found that there is no general rule that vocal-tract information contributes more to perceived speaker identity than voice-source information. Sometimes vocal-tract information is more important, sometimes voice-source information. The results fit the subjective cues that listeners reported they had been using to perform the speaker identification task quite well. Besides the expected importance of the vocal-tract filter, the spectral balance between high- and low-frequency components of the voice-source spectrum and the flutter of F_0 proved to be important perceptual cues for speaker identity.

SATURDAY MORNING, 31 OCTOBER 1992 EXPLORER'S ROOM, 7:55 A.M. TO 12:30 P.M.

Session 1aUW

Underwater Acoustics: Seismo-Acoustics and Ambient Noise

Joseph F. Gettrust, Chair
Naval Research Laboratory Detachment, Code 362, Stennis Space Center, Mississippi 39529-5004

Chair's Introduction—7:55

Contributed Papers

8:00

1aUW1. Experimental verification of the theory for sound propagation in a wedge with a shear supporting bottom. Stewart A. L. Glegg (Ctr. for Acoust. and Vib., Dept. of Ocean Eng., Florida Atlantic Univ., Boca Raton, FL 33431) and Grant B. Deane (Scripps Inst. of Oceanog., La Jolla, CA 92093)

Laboratory experiments have been carried out to investigate underwater sound propagation over a sloping bottom that can support shear. The sound field was measured in both the across slope and down slope direction for many different frequencies and bottom slopes. The experimental results have been compared with the source image theory for sound propagation in a penetrable wedge [G. B. Deane, J. Acoust. Soc. Am. **91**, 2390 (A) (1992)] and have shown very good agreement for both down slope and across slope propagation. In particular the effect of shear waves on across slope propagation has been demonstrated both theoretically and experimentally. Further the effect of a sediment layer has been successfully incorporated into the source image theory for