Tilburg University

Categorical perception of emotional speech

de Gelder, B.; Vroomen, J.

Published in:
Journal of the Acoustical Society of America

Publication date:
1996

9:00

5aPP5. Is impaired auditory temporal processing a cause of speech-language disorders? Negative evidence from psychoacoustic investigations. Charles S. Watson and Betty U. Watson (Dept. of Speech and Hearing Sci., Indiana Univ., Bloomington, IN 47405)

Psychoacoustic studies cast some doubt on the hypothesis that deficits in auditory temporal processing cause speech and language disorders. First, the temporal acuity of the auditory system so greatly exceeds that required by tasks considered to support this hypothesis [e.g., Tallal and Piercy, Nature **241**, 468–469 (1973)] that disordered subjects would be quite unable to understand speech if their basic temporal processing were as imprecise as it is proposed to be. Second, in a study of 127 normal-hearing listeners, nonsense-syllable recognition was not correlated with temporal discrimination measured with various nonspeech stimuli, and the same result was later obtained for sentences, words and CV's, although all speech measures load on a common factor. Third, structural-equation models were fitted to data collected from 24 college students with specific reading disabilities, and 70 who were "normal." Data to be predicted included passage comprehension, word-attack skills, and word identification. A model using measures of speech perception and of intelligence yielded $0.7>r^2>0.9$, but these predictions were not improved by nonspeech measures of auditory temporal processing. A viable hypothesis is that sensory processing speed may be only indirectly related to some speech-language disabilities, but more directly related to intelligence. [Work supported by NIH/NIDCD.]

9:15

5aPP6. Phoneme recognition by a non-native English-speaking cochlear implantee using the WSP, MSP, and spectra speech processors: A case study. Kathryn S. Copmann (Dept. of Speech Lang. Pathol. /Audiol., Loyola College, 4501 N. Charles St., Baltimore, MD 21210-2699)

This study presents the performance of a non-native English-speaking adult using the Nucleus 22 Channel cochlear implant. Included are the subject's (1) history, (2) promontory stimulation results, (3) audiological information, (4) pre- and post-implantation speech recognition test results, (5) post-implantation auditory phoneme recognition using the wearable speech processor (WSP), the minispeech processor (MSP), and the Spectra processor, and (6) the subject's subjective reports of benefits in terms of communication and social interactions. The subject, a highly motivated and intelligent male who, presented with a profound bilateral, progressive sensorineural hearing loss of unknown etiology, was 59 years old when the right ear was implanted. Prior to implantation the subject was aided in the left ear, but had never been aided in the right ear. The subject's performance using each of three processors is presented for medial vowel recognition and medial consonant recognition using auditory stimulation only. Results are displayed through confusion matrices generated based on the subject's responses to these acoustic stimuli. In summary, the vowel and consonant phoneme recognition of a subject with the Nucleus cochlear implant using various processors is presented. Improved phoneme recognition is demonstrated with filter banking versus feature extraction coding strategies.

9:30

5aPP7. A handicapped listener's articulation tests underwater. Akiteru Fukuda and Heiji Okada (Dept. of Commun., Shibaura Inst. of Tech., 3-9-14 Minato-ku Tokyo, 108 Japan)

There have been a number of papers in underwater sound propagation for a diver's mutual communication. A healthy man's monosyllable hearing test underwater has been reported previously. Now, a report of results on disabled (hard of hearing) listeners test is given. Sound propagation tests used a standard Japanese 25 m(L)×13 m(W)×1.5 m(D) waterpool. An underwater speaker Altec type UW-30 is used and is driven by a 10-W amplifier. A listener's monosyllable test signal is used with a standard voice number (101) by an NHK announcer. The distances between the radiation speaker and the disabled listener are agreed upon at 5, 10, and 20 m, and depths are changed from 1.3 to ~1.5 m. Radiator and listener's depths are held at the same level through every test. Depth simulation used with the air chamber at 10 kg/cm^2 normal air at Saitama Medical Univ., where pressures are controlled equivalent to 0-, 10-, 20-, and 30-m water depths. In the results of the articulation test, about the same scores are observed between the handicapped person and the healthy listener. Especially at the distance of 10-m, a handicapped person's average score is greater than 4% compared with the healthy person's score. Such a result requires further examinations.

9:45

5aPP8. Learning and generalization in auditory backward masking. Beverly A. Wright, Paul A. Johnston, and Miriam D. Reid (Keck Ctr., Box 0732, Univ. of California, San Francisco, CA 94143-0732)

Excessive auditory backward masking has been observed in elderly persons and in children with language impairments. The purposes of this project were to determine (1) whether practice can improve the ability to detect a backward-masked signal and, if so, (2) whether this improvement generalizes to other masking tasks. These issues were examined in six normal-hearing adults using an adaptive, 2IFC procedure. Each subject was asked to detect a tonal signal (10 ms, 1000 Hz) presented immediately before a noise masker (300 ms, 200–1800 Hz, 40-dB SPL spectrum level), on 900 trials per day for ten days. Before and after the training period, subjects were tested on five other masking tasks in addition to the trained task. A one-way ANOVA revealed a significant learning curve over the ten days of training [$F(9,45)=5.1$, $p<0.001$], with the average signal threshold decreasing from 63 to 53 dB SPL. Comparison of the pre- and post-test results showed significant threshold reductions in other backward-masking tasks, but not in simultaneous- and forward-masking tasks. These results indicate that (1) the interference from an auditory backward masker can be reduced with practice, and that (2) this learning is specific to backward masking. [Work supported by the James S. McDonnell Foundation and NIDCD.]

10:00–10:15 Break

10:15

5aPP9. Categorical perception of emotional speech. Bea de Gelder and Jean Vroomen (Dept. of Psych., Tilburg Univ., Tilburg, The Netherlands)

Emotions are communicated through the prosody of speech. Are such expressions perceived categorically or rather in a multidimensional space without clear foci? The former alternative is in line with longstanding views about the existence of some basic emotions. Recent reports present evidence that facial expressions are perceived categorically [de Gelder *et al.*, Cognit. Emotion (in press)]. To explore this issue for auditory speech, a continuum between happiness and fear was created by varying pitch excursion, pitch height, and duration of an utterance via PSOLA. Categoricity of perception for auditory expressions was examined in (a) an auditory-only presentation and (b) an auditory–visual condition of simultaneous presentation of the auditory stimuli and facial expressions corresponding to the either of the two expressions. In the auditory-only conditions a clear CP phenomenon is observed. In the bimodal conditions the visual information modulates the perception of the speech emotion. The latter phenomenon is not observed in a patient suffering from prosopagnosia. The findings raise the issue of a common biological basis for facial and vocal expression perception.

10:30

5aPP10. Phoneme detection in resyllabified words. Jean Vroomen and Bea de Gelder (Dept. of Psych., Tilburg Univ., Tilburg, The Netherlands)

There is evidence that the human listener takes, in continuous speech, the onset of a syllable as the onset of a spoken word [J. Vroomen and B. de Gelder, J. Exp. Psychol.: Human Percept. Performance (in press)]. However, this syllabic strategy would fail in resyllabified words like "my bike is," pronounced as /mai bai kis/. In the present study, a phoneme