Hearing Aid Research

RLE Groups
Sensory Communication Group, Auditory Perception and Cognition Group

Sponsor
National Institutes of Health Grant R01 DC00117

Academic and Research Staff
Professor Louis D. Braida, Dr. Julie Greenberg, Dr. Jean C. Krause, Dr. Andrew Oxenham, Dr. Karen L. Payton

Visiting Scientista and Research Affiliates
Dr. Paul Duchnowski, Dr. Kenneth W. Grant, Dr. Christine Koh, Dr. Christine M. Rankovic, Dr. Peninah S. Rosengard

Graduate Students
Shaoyan Chen, Raymond Goldsworthy, Xueyan Li, David P. Messing, P. Ramanand Nayak

Technical And Support Staff
Loraine Delhorne, Joseph A. Frisbie, Denise Rosetti

Our long-term goal is to develop improved hearing aids for individuals who suffer from sensorineural hearing impairments and cochlear implants for the deaf. Our efforts are focused on problems resulting from inadequate knowledge of the effects of various transformations of speech signals on speech reception by impaired listeners, specifically on the fundamental limitations on the improvements in speech reception that can be achieved by processing speech. Our aims are

To evaluate the effects of style of speech articulation and variability in speech production on speech reception by hearing impaired listeners.

To develop and evaluate analytical models that can predict the effects of a variety of alterations of the speech signal on intelligibility.

To develop and evaluate signal processing techniques that hold promise for increasing the effectiveness of hearing aids.

To assess the relative contributions of various functional characteristics of hearing impairments to reduced speech-reception capacity.

Studies and Results

Characteristics of the Speech Signal

Three traditional ASR parameterizations matched with Hidden Markov Models (HMMs) were compared (Sroka and Braida, 2005) to humans for speaker-dependent consonant recognition using nonsense syllables degraded by highpass filtering, lowpass filtering, or additive noise. Confusion matrices were determined by recognizing the syllables using different ASR front ends, including Mel-Filter Bank (MFB) energies, Mel-Filtered Cepstral Coefficients (MFCCs), and the Ensemble Interval Histogram (EIH). In general the MFB recognition accuracy was slightly higher than the MFCC, which was higher than the EIH. For syllables degraded by lowpass and highpass filtering, automated systems trained on the degraded condition recognized the consonants as well as humans. For syllables degraded by additive speech-shaped noise, none of the automated systems recognized consonants as well as humans. The greatest advantage displayed by humans was in determining the correct voiced/unvoiced classification of consonants in noise.

Computational Model of Speech Intelligibility

We reported a comparison of the predictions of models of integration to data on the reception of consonants filtered into a variety of frequency bands (Ronan et al, 2005). New data on the consonant identification were presented. Three experiments were conducted testing the following bands: Experiment I, 0-2100 Hz and 2100-4500 Hz; Experiment II, 0-700 Hz combined with 700--1400, 1400-2100, 2100-2800, and 2800-4500 Hz; Experiment III all combinations of 700--1400, 1400-2100, 2100-2800, and 2800-4500 Hz. The predictions of four models, Fletcher's (1950) independent errors model, Massaro's Fuzzy Logical Model of Perception (1987), and Braida's Pre-Labelling and Post-Labelling Models of Integration (1991) were compared in terms of their ability to predict combined-band scores. At least two models were capable of predicting performance for each combined-band condition. For Experiment I, all models were able to make satisfactory predictions. For Experiment II, a variant of the Pre-Labelling model was able to make satisfactory predictions. For Experiment III, no model was able to make satisfactory predictions, but the Fuzzy Logical Model of Perception and a variant of the Pre-Labelling model made relatively good predictions. Thus the ability of the models to predict performance depended more on whether the condition included the lowest frequency band than on the adjacency or frequency separation.

In the area of STI research, Xueyan Li, is completing a thesis project on: Statistical Analysis of Speech Envelope Spectra. For this project, Xueyan is analyzing the statistical properties of speech envelope spectra for different speaking styles (clear/slow, clear/norm and conversational/norm) and multiple talkers. The goal is to identify the number of sentences needed to reliably compute an envelope spectrum and to generate confidence intervals that will show the extent to which the speaking styles are statistically distinct.

Shaoyan Chen, a Ph.D. student working on her dissertation proposal, has recently completed preliminary work to provide more consistent STI values for amplitude-compressed speech conditions. She has determined that a modulation reduction factor can be analytically determined based on compression ratio and SNR. If the STI is compensated by the modulation reduction factor, the amplitude compression and linear amplification intelligibility data reported by Rosengard et al. have consistent STI values.

Cochlear-Implant Research

One component of our work in the area of models of speech intelligibility concerns predicting the intelligibility of cochlear-implant processed speech. The goal is to develop a subject-independent metric that can be used to predict the maximum possible intelligibility performance for a particular cochlear-implant speech processing strategy. (Subject-dependent factors may lead to lower performance for particular subjects.) This metric will then be used to evaluate promising noise-reduction strategies for cochlear implant preprocessing. In this effort, our previous experience with hearing-aid users in two main areas (models of speech intelligibility and signal processing algorithms for noise reduction) is being applied to benefit cochlear-implant users, a population for whom background noise affects speech intelligibility even more adversely than hearing-aid users.

The model of speech intelligibility under consideration is based on the Speech Transmission Index (STI). STI was originally developed as a way of assessing room acoustics, and the original STI calculations are based on a system's response to specific test signals. Although these test signals are appropriate for assessing room acoustics, they are not suitable for many kinds of signal processing used in hearing aids and cochlear implants. As a result, several research groups, including our own, have attempted to develop methods for calculating STI based on the speech signal itself, rather than specific test signals. We have analyzed the various methods for speech-based STI calculation described in the literature, established explicit relationships between the various speech-based STI calculations, identified several problems with existing methods, and proposed improved methods for speech-based STI calculations (Goldsworthy and Greenberg, 2004).

We have performed a set of four intelligibility experiments aimed at 1) understanding the factors that affect the intelligibility of cochlear-implant processed speech; and 2) performing a comprehensive evaluation of the various methods for calculating speech-based STI. Using normal-hearing and cochlear-implant listeners as subjects, the experiments address the following conditions: acoustic degradations (reverberation and additive noise), N of M processing (used in commercial cochlear-implant speech processors), spectral subtraction, and binaural noise reduction. Our results show that: 1) both spectral subtraction and binaural noise reduction improve the intelligibility of cochlear-implant processed speech; and 2) one of the proposed speech-based STI methods consistently predicts the major trends in speech reception scores for all four experiments.

Signal Processing for Hearing Aids

An MS student, P. Ramanand Nayak, is completing an MS thesis entitled: Amplitude Compression of Speech Using Wavelet Transform and Performance Comparison of Different Wavelet Families. Ramanand implemented multi-level compression of 4 different wavelet families at three different compression ratios (2, 3 and 5). Based on informal qualitative listening tests, the Coiflet wavelet had the highest quality over all. The wavelet results are also going to be compared to a traditional 4-band filter implementation of amplitude compression.

Characteristics of Sensorineural Hearing Impairment

A loss of cochlear compression may underlie many of the difficulties experienced by hearing-impaired listeners. Two behavioral forward-masking paradigms that have been used to estimate the magnitude of cochlear compression are growth of masking (GOM) and temporal masking (TM). The aim of this study (Rosengard et al., 2005) was to determine whether these two measures produce within-subjects results that are consistent across a range of signal frequencies and, if so, to compare them in terms of reliability or efficiency. GOM and TM functions were measured in a group of five normal-hearing and five hearing-impaired listeners at signal frequencies of 1000, 2000, and 4000 Hz. Compression values were derived from the masking data and confidence intervals were constructed around these estimates. Both measures produced comparable estimates of compression, but both measures have distinct advantages and disadvantages, so that the more appropriate measure depends on factors such as the frequency region of interest and the degree of hearing loss. Because of the long testing times needed, neither measure is suitable for clinical use in its current form.

In the area of hearing loss simulations, Rashmi Prabhakar completed an MS thesis in January 2005 entitled: Implementation of a Hearing Loss Simulation on DSP Hardware. For this thesis, Rashmi implemented the sloping and flat hearing-loss simulations used by Rosengard et al. (in press) on an Analog Devices Sharc DSP board. These simulations included recruitment. While a prototype, the simulation confirmed the feasibility of real-time simulations of specific hearing loss profiles.

PUBLICATIONS

Journal Articles

Published

Goldsworthy, R. and Greenberg, J. E. (2004). “Analysis of speech-based Speech Transmission Index methods with implications for non-linear operations,’’ J. Acoust. Soc. Am., **116** (6), 3679-3689.

Krause, J.C. and Braida, L.D. (2004). "Acoustic properties of naturally produced clear speech at normal speaking rates", J. Acoust. Soc. Am. **115** (1), 362-378.
Ronan, D, Dix, A, Shah, P. and Braida, L.D. (2004). "Integration of acoustic cues for consonant Identification across frequency bands", J. Acoust. Soc. Am. **116** (3), 1749-1762.

Rosengard, P.S., Oxenham, A. J., and Braida, L.D. (2005). "Comparing different estimates of cochlear compression in listeners with normal and impaired hearing." J. Acoust. Soc. Am. **117** (5) 3028-3041.

Sroka, J.J., and Braida. "L.D. (2005). Human and machine consonant recognition," Speech Communication **45**, 401-423.

Accepted for Publication

Rosengard, P.S., Payton, K.L., and Braida, L.D. (2005). "Effects of slow-acting wide dynamic range compression on measures of intelligibility and ratings of speech quality in simulated-loss listeners," accepted for publication in J. Speech and Hearing Res.

Submitted for Publication

Chen, Shaoyan and Payton, Karen L., "Compensating the STI to Predict Intelligibility for Amplitude Compressed Speech", Submitted to IEEE Workshop App. Sig. Proc. Audio Acoust., 2005.

Chapters in Books

Reed, C.M. and Braida, L.D. (2004). "Frequency Compression,'' in the M.I.T. Encyclopedia of Communication Disorders, R. Kent Ed., The M.I.T. Press, Cambridge, MA.. 471-475.

Meeting Papers Presented

Li, Xueyan and Payton, Karen, "Statistical Analysis of Speech Envelope Spectra", 2005 Sigma Xi Research Exhibition, April 26, 27, 2005.

Nayak, P. Ramanand and Payton, Karen, "Amplitude Compression of Speech Using Wavelet Transform and Performance Comparison of Different Wavelet Families", 2005 Sigma Xi Research Exhibition, April 26, 27, 2005.

Theses

Goldsworthy, Raymond, "Improving Speech Intelligibility for Cochlear Implant Users," MIT PhD thesis, Jan. 2005.

Prabhakar, Rashmi, "Implementation of a Hearing Loss Simulation on DSP Hardware", MS Thesis, Univ. Mass. Dartmouth, January 2005.

Nayak, Pilar Ramanand, "Amplitude Compression of Speech Using Wavelet Transform and Performance Comparison of Different Wavelet Families", MS Thesis, Univ. Mass. Dartmouth, June 2005.