Hearing Aid Research

Sponsor
National Institutes of Health Grants R01 DC00117, R01 DC007152, AFOSR Contract No. FA9550-05-C-0032.

Academic and Research Staff
Professor Louis D. Braida, Dr. Joseph Desloge, Dr. Raymond Goldsworthy, Dr. Karen L. Payton, Dr. Charlotte M. Reed

Visiting Scientists and Research Affiliates
Dr. Paul Duchnowski, Dr. Oded Ghitza, Dr. Kenneth W. Grant, Professor Ying-Yee Kong,
Professor Jean C. Krause, Dr. Peninah S. Rosengard

Technical and Support Staff
Lorraine Delhorne, Denise Stewart

Our long-term goal is to develop improved hearing aids for people suffering 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 develop and evaluate analytical models that can predict the effects of a variety of alterations of the speech signal on intelligibility.

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

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

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

Studies and Results

I-A. Role of Audibility in Speech and Psychoacoustic Performance of Listeners with Cochlear Hearing Loss

This research is concerned with analyzing the factors responsible for poor speech reception by listeners with hearing impairments, and with developing techniques for overcoming these degradations. To the extent the research is successful, it will help determine design goals for improved wearable hearing aids, establish new criteria and techniques for aural rehabilitation, and contribute to improved understand of both residual auditory function and speech perception.

Research during the past year has been directed at examining the role of audibility in explaining the ability of listeners with moderate-to-profound sensorineural hearing loss to understand speech in noise. Individual hearing losses are simulated in age-matched normal-hearing listeners using masking noise and multi-band expansion to model the effects of cochlear hearing loss. Stimuli presented to hearing-impaired and simulated-loss listeners are thus equated in stimulus level specified in both dB SPL and dB SL. Any differences in performance observed on speech-reception tests between hearing-impaired and normal-hearing listeners can then be ascribed to supra-threshold deficits associated with hearing impairment. A battery of psychoacoustic measurements is employed to determine the source of any such suprathreshold components to speech-reception performance.

Progress over the past year is reported in three major areas: (1) experiments examining speech reception in noise for listeners with real and simulated hearing impairment; (2) experiments examining the spectral, temporal, and cognitive abilities of listeners with real and simulated hearing impairment; and (3) a critical review of the literature on the role of audibility in the temporal, intensive, and spectral abilities of listeners with cochlear hearing loss.

I-B-1. Speech Reception in Noise for Listeners with Real and Simulated Hearing Impairment.

This study examined the effects of audibility and age on the release from masking (RM) for speech in interrupted versus steady-state noise in listeners with real and simulated hearing loss. The absolute thresholds of each of ten hearing-impaired (HI) listeners with bilaterally symmetrical losses were simulated in normal-hearing listeners through a combination of spectrally-shaped masking noise and multi-band expansion for the octave bands with center frequencies from 0.25-8 kHz. Each of the ten individual hearing losses (which included flat, sloping high-frequency, and inverted cookie-bite configurations) was simulated in two groups of three different normal-hearing listeners (one age-matched group and one non-age-matched group, resulting in a total of 60 normal-hearing subjects). The speech-to-noise ratio (S/B) for 50%-correct identification of HINT sentences (Nilsson et al., 1994) was measured in backgrounds of steady-state noise and temporally-modulated (10 Hz square-wave) noise. Interrupted and continuous background noise was presented at two overall levels (65 and 80 dB SPL). In addition, speech-reception thresholds were obtained in a continuous background noise of 30 dB SPL to approximate SRT. Speech was presented using three different hearing-aid configurations: (a) unprocessed (i.e., no hearing aid), (b) linear hearing aid using the NAL-RP prescription (Dillon, 2001), and (c) a compressive hearing aid based upon the algorithms developed by Goldstein (Goldstein et al., 2003).

Measurements of S/B for each hearing-impaired listener at each of the 15 listening conditions (5 noise conditions X 3 hearing-aid configurations) were compared to those obtained in the age-matched and non-age-matched groups of normal-hearing listeners with simulated loss. Observed RM values (i.e., the difference in S/B obtained in steady-state minus interrupted noise) and the significance of the RM values from 0 dB appear to be related to the audibility of the noise, which is directly affected by the degree of hearing impairment as well as the use of NAL amplification. The Speech Intelligibility Index (SII; ANSI S3.5-1997) and an extension of the SII to interrupted noise (Rhebergen and Versfeld, 2005) were used to model the performance of the hearing-impaired listeners. Predictions of RM derived from the SII provide a good match to the observed values for individuals with real and simulated hearing loss. Our results indicate that the masking release appears to be determined primarily by audibility regardless of subject age.

I-B-2. Measures of Spectral, Temporal, and Cognitive Processing in Listeners with Real and Simulated Hearing Impairment

Five experiments have been conducted on the ten listeners with cochlear hearing impairment who participated in the speech-in-noise study described above in Sec. B-2. Each hearing loss was simulated in a group of three normal-hearing listeners matched roughly in age to a given hearing- impaired individual. Psychoacoustic tests were selected to probe aspects of spectral and temporal resolution thought to be important for the reception of speech in noise. A measure of cognitive function was included in the testing.

Spectral resolution is examined using notched-noise masking. The detection of 200-msec probe tones (250, 500, 1000, 2000, and 4000 Hz) was examined in a 220-ms notched-noise simultaneous masker consisting of two bands of noise (one located above and the other below the probe signal, each with bandwidth of 0.25 times the frequency of the probe tone). The bands of noise were selected to create notch widths ranging from 0 to 0.8 times the frequency of the probe tone. Masked thresholds as a function of notch width were analyzed using a roex filter model (Patterson and Nimmo-Smith, 1980). Comparisons of resulting estimates of the equivalent rectangular bandwidth at each test frequency will be made between listeners with real and simulated hearing loss.

Aspects of temporal resolution were examined using three different techniques. (a) In tone-on-tone forward masking, the detection of a 10-msec probe (500, 1000, 2000, and 4000 Hz) is examined using 110-ms maskers (at frequencies of 0.55, 1.0 and 1.15 times the probe frequency) and five values of delay time between offset of masker and onset of probe (0, 10, 20, 60, and 100 ms). Slopes of the masking functions were calculated for each masker frequency and compared across listeners with real and simulating hearing loss. (b) In a temporal-modulation detection task, the ability to discriminate a continuous 500-ms broadband noise from an amplitude-modulated noise was examined at each of ten modulation frequencies (ranging from 2 to 1024 Hz). The resulting temporal-modulation transfer functions for individual HI and simulated-loss listener were analyzed using an exponential fitting procedure to estimate the time constant and the interpolated DC values of the functions. (c) Temporal gap detection is measured for auditory and tactual stimuli using 250 and 400 Hz leading and trailing markers with a nominal duration of 100 ms and a reference gap of 6.4 ms. Results are being analyzed in terms of spectral-disparity effects of leading and trailing markers and the relationship between thresholds obtained for auditory versus tactile presentation. Comparisons are made between listeners with real and simulated hearing loss.

Finally, a test of cognitive ability is included in the test battery. Subjects are administered a Reading-Span test (Ronnberg et al.,1999) that examines the ability to retain the final word of orally- read sentences in lists of nonsense and sensible sentences ranging in set size from 2 to 5. This cognitive measure will be related to performance on the speech and psychoacoustic tests.

I-B-3. Review of Past Research on the Role of Audibility in Predicting Effects of Hearing Impairment

Although supra-threshold effects of hearing impairment are widely believed to be related to the decreased resolution on psychoacoustic tasks and the poorer speech-reception abilities of hearing- impaired listeners, the role of reduced audibility itself in explaining the consequences of hearing loss is as yet not completely understood. We have completed a critical review of research on temporal resolution in listeners with hearing impairment (Reed, Braida, and Zurek, 2009). We are currently working on two manuscripts concerning the role of audibility on (a) intensity and loudness perception and (b) spectral resolution in hearing-impaired listeners.

I-1-C. Significance

Our research is concerned with analyzing the factors responsible for poor speech reception by listeners with hearing impairments, and with developing techniques for overcoming these degradations. To the extent the research is successful, it will help determine design goals for improved wearable hearing aids, establish new criteria and techniques for aural rehabilitation, and contribute to improved understand of both residual auditory function and speech perception.

I-D Plans for the Coming Year

Area 1: Speech Reception in Noise. We will complete the analysis of data collected on 10 HI listeners and a total of 60 NH listeners with simulated hearing-loss and prepare a manuscript for publication.

Area 2: Spectral, Temporal, and Cognitive Tests: We will complete the analysis of data collected on 10 HI subjects and 20 NH subjects with simulated loss; complete data collection and analysis on an additional 10 NH subjects; and prepare manuscripts for publication on the topics of notched-noise masking, forward masking, and temporal-modulation detection.

Area 3: Review of the Literature on the Role of Audibility in Cochlear Hearing Loss: We will complete the preparation of manuscripts in the areas of intensity and loudness perception and spectral resolution; and begin work on a manuscript concerned with speech reception.

II-A. Models of Speech Intelligibility

This research is directed at developing and experimentally evaluating models of speech intelligibility for impaired listeners, that is, robust metrics that predict speech reception scores for a variety of acoustic degradations and speech processing conditions. To this end we have four aims:

Aim 1) Measure speech reception in three classes of listeners (moderately-to-severely hearing impaired, cochlear-implant users, and normal-hearing subjects listening through a channel-vocoder simulation of cochlear-implant sound processing) for four types of alterations of speech (acoustic degradations arising from noise and reverberation, band-pass filtering, amplitude compression, and noise-reduction algorithms).

Aim 2) Characterize the basic abilities of hearing-impaired and CI listeners (in terms of basic sensitivity, dynamic range, spectral resolution, and temporal resolution) and their ability to integrate information across different filtered bands of speech.

Aim 3) Develop STI-based metrics of speech intelligibility and apply these metrics to the stimuli used to test hearing-impaired and CI listeners in AIM1 above. The metrics will incorporate the individual listener characteristics obtained in (2).

Aim 4) Evaluate the metrics developed in (3) by comparing metric predictions for a variety of listeners and speech processing conditions to the data obtained in (1).

II-B Cochlear Implant Research

To date 14 implantees have been tested, including 5 bilateral implantees. The psychoacoustic measures include: pure tone intensity discrimination, pure tone and tone complex frequency discrimination, forward and backward masking of a 300ms tone burst upon a 10ms tone burst, gap detection within noise bands of various bandwidths, tone-on-tone and synthetic formant-on-formant masking, tone-in-noise detection, and amplitude modulation detection. Psychoacoustic measures were determined using either two- or three-alternative forced choice (AFC) paradigms. The speech reception measures include vowel and consonant recognition in quiet, in speech-shaped noise, and in gated speech-shaped noise. Noise levels were adjusted adaptively to estimate the speech reception threshold (SRT). The results for these measures are being compared with normal hearing performance in order to understand differences between cochlear implantees and normal hearing listeners. The relative performance of implantees and normal hearing listeners indicates that implantees show large deficits in tone-on-tone masking (20-40 dB), forward and backward masking (10 dB), and tone-in-noise detection (10-15 dB). This research conducted at MIT was presented at the 2009 conference for the Association for Research in Otolaryngology (***Goldsworthy *et al.*, 2009).

The implantees were tested on 20 measures of perception (14 psychoacoustic and 6 phoneme perception measures). We found that measures of temporal and spectral resolution using simple tone detection tasks are excellent predictors of phoneme perception both in quiet and in noise. Strong correlation was also found between consonant perception in quiet and backward masking (*e.g.*, Figure 1). But more important than individual correlations between measures, is the depth of analysis that this approach provides. An important goal of the grant is to integrate such psychoacoustic assessment into the modeling framework.

In addition to these measures that use acoustic stimuli that are processed through recipients' clinical sound processors, we have also begun pilot studies in which electrical stimuli are specified and delivered to the subject's electrode array using the Laura system, bypassing the implant sound processor. These pilot studies are examining temporal rate pitch perception, forward and backward masking of different electrode pairs, and simultaneous masking of different electrode pairs.

As an example, one pilot study that we are undertaking evaluates the difference limens of rate-pitch for electric pulse trains delivered to a single electrode. The reference stimulus is a 400ms pulse train with a rate specified as the condition variable. Rates tested in the pilot were 55, 110, 220, and 440Hz. The target stimulus has a rate that is higher than the reference, with a starting value in an adaptive test that is twelve semitones greater than the reference. The levels of the pulse train are randomized between 60 and 100% of the subject's electric dynamic range. The task is a 3AFC paradigm where the subjects are instructed to select the interval that sounds different. An initial result is illustrated in Figure 2. The subject was capable of performing the task, with no training, with a resolution of approximately 3 semitones (roughly 19%) for the 55, 110, and 220Hz frequencies, but could not perform the task at 440Hz. Our results are in agreement with the literature (*e.g.*, Eddington *et al.*, 1978; Shannon, 1983; Tong *et al.*, 1982; Townshend *et al.*, 1987; Carlyon *et al.*, 2008) suggesting a general failure to convey fine timing cues above 300Hz.

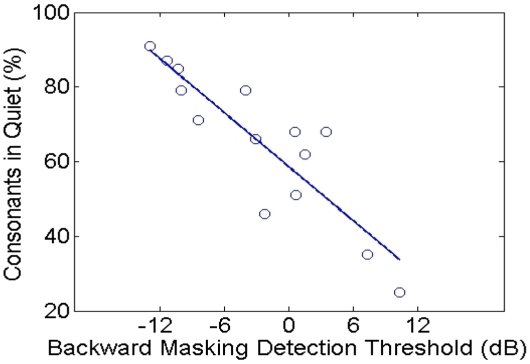

Figure 1. An example relation between psychoacoustic measurement (backward masking) and speech reception (identification of consonants in quiet).

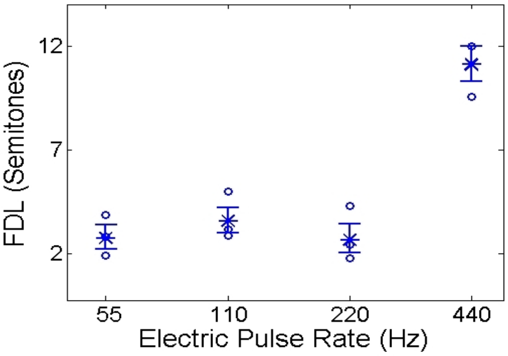

Figure 2. Results of a laboratory processor controlled measurement of the relation between the electrical pulse rate and the difference limen for pulse rate.

II-C STI Based Models of Speech Intelligibility

We presented the progress reported last year at two international conferences: Acoustics’08 and the 9th Congress of the International Commission on the Biological Effects of Noise (ICBEN). A 4-page paper was published in each proceeding (Payton and Shrestha 2008a, 2008b). The paper in the Acoustics’08 proceedings (Payton and Shrestha 2008a) focused on the short-time STI prediction results for words spoken clearly at normal rates and conversationally. The paper published in the ICBEN proceedings (Payton and Shrestha 2008b) focused on the short-time STI results for conversational speech in speech-shaped Gaussian noise and in multi-talker babble noise. Currently the data from both papers is being consolidated into a single manuscript for submission to J. Acoust. Soc. Am. later this year.

Publications

Krause, J. C. and Braida, L. D. "Evaluating the role of spectral and envelope characteristics in the intelligibility advantage of clear speech," J. Acoust. Soc. Am., 125 (5), 3346–3357, 2009.

Messing, D. P., Delhorne, L., Bruckert, E., Braida, L D., and Ghitza, O. “A non-linear efferent-inspired model of the auditory system: matching human confusions in stationary noise, “ Speech Communication 51, 668-683, 2009.

Reed, C.M., Braida, L.D., and Zurek, Patrick M. (2008). "Review of the literature on temporal resolution in listeners with cochlear hearing impairment: A critical assessment of the role of suprathreshold deficits," Trends in Amplification, 12, No. 5, in press (for Winter 2008 issue).

Meeting Papers

Goldsworthy, R., Delhorne, L., and Braida, L. (2009). "Relations between Psychoacoustic and Speech Reception Measures in Cochlear Implant Users," Poster presented at 32nd Midwinter Research Meeting of the Association for Research on Otolaryngology, Feb. 19-24, 2009, Baltimore, MD.

Payton, K. L., Shrestha, M. (2008a). "Analysis of short-time Speech Transmission Index algorithms," Proc. Acoustics'08, June 29-July 4, 2008, Paris, France

Payton, K. L., Shrestha, M. (2008b). "Evaluation of short-time speech-based intelligibility metrics," Proc. Int. Commission on Biol. Effects Noise 2008, July 21-25, 2008, Foxwoods, CT.

References:

ANSI (1997). "ANSI S3.5-1997, American national standard methods for calculation of the speech intelligibility index" (American National Standards Institute, New York).

Carlyon R. P., Long C. J., and Deeks J. M. (2008). "Pulse-rate discrimination by cochlear-implant and normal-hearing listeners with and without binaural cues," *JASA* 123(3), 2276-2286.

Dillon, H. (2001). *Hearing Aids*, Thieme, New York.

Eddington DK, Dobelle WH, Brackmann DE, Mladejovsky MG and Parkin JL (1978). "Place and periodicity pitch by stimulation of multiple scala tympani electrodes in deaf volunteers," In: *Trans Am Soc Art Intern Organs* 24: 1-5.

Goldstein, J.L., Oz, M., Gilchrist, P.M., and Valente, M. (2003). Signal processing strategies and clinical outcomes for gain and waveform compression in hearing aids. Proc. 37th Asilomar Conf. on Signals, Systems, and Computers, 391-398 (IEEE Pub. ISBN 0-7803-8104-1.

Nilsson, M., Soli, S.D., and Sullivan, J.A. (1994). "Development of the Hearing in Noise Test for the measurement of speech reception thresholds in quiet and in noise," J. Acoust. Soc. Am., 95, 1085-1099.

Patterson, R.D., and Nimmo-Smith, I. (1980). "Off-frequency listening and auditory-filter asymmetry," J. Acoust. Soc. Am., 67, 229-245.

Payton, K. L. and Shrestha, M. (2008a) "Evaluation of short-time speech-based intelligibility metrics," in *Proc. 9th Cong. Int. Commiss. Biol. Effects Noise*, July 2008, 243-251.

Payton, K. L. and Shrestha, M. (2008b) "Analysis of short-time Speech Transmission Index algorithms", in *Proc. Acoustics '08*, June 2008, 633-638.

Reed, C.M., Braida, L.D., and Zurek, Patrick M. (2009). "Review of the literature on temporal resolution in listeners with cochlear hearing impairment: A critical assessment of the role of suprathreshold deficits," Trends in Amplification, 13, 4-43.

Rhebergen, K.S., and Versfeld, N.J. (2005). "A Speech-Intelligibility Index-based approach to predict the speech threshold for sentences in fluctuating noise for normal-hearing listeners," J. Acoust. Soc. Am., 117, 2181-2192.

Ronnberg, J., Andersson, J., Samuelsson, S., Soderfeldt, B., Lyxell, B., and Risberg, J. (1999). “A speechreading expert: The case of MM,” J. Speech Hearing Lang. Res., 42, 5-20.

Shannon R. V. (1983). “Multichannel electrical stimulation of the auditory nerve in man. I. Basic psychophysics,” Hear Res 11, 157-189.

Tong Y., Clark G.M., Blamey P., Busby P., Dowell R. (1982). “Psychophysical studies for two multiple- channel cochlear implant patients,” *JASA* 71, 153-160.

Townshend B., Cotter N., van Compernolle D. and White R.L. (1987). “Pitch perception by cochlear implant subjects,” *JASA* 82, 106-115.

Byrne, D. and Dillon, H. (1986). “The National Acoustics Laboratory new procedure for selecting the gain and frequency response of a hearing aid,” Ear and Hearing 7, 257-265.

Dolan, D. F., and Nuttal, A. L. (1988). “Masked cochlear whole-nerve response intensity functions altered by electrical stimulation of the crossed olivocochlear bundle,” J. Acoust. Soc. Am. 83: 1081–1086.

Goldsworthy, R. and J. Greenberg, (2004). “Analysis of speech-based speech transmission index methods with implications for nonlinear operations,” J. Acoust Soc Am. 116(6), 3679-3689.

Holube, I. and B. Kollmeier (1996). “Speech intelligibility prediction in hearing-impaired listeners based on a psychoacoustically motivated perception model.” J. Acoust. Soc. Am. 100(3): 1703-1716.

T. Houtgast, H. J. M. Steeneken, and R. Plomp (1980), "Predicting Speech Intelligibility in Rooms from the Modulation Transfer Function I. General Room Acoustics," Acustica, 46, 60-72.

Koch, R. (1992). Gehörgerechte Schallanalyse zur Vorhersage und Verbesserung der Sprachverständlichkeit (Auditory sound analysis for the prediction and improvement of speech intelligibility), Universität Göttingen.

Krause, J. C. (2001), "Properties of Naturally Produced Clear Speech at Normal Rates and Implications for Intelligibility Enhancement," Ph.D., Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA, USA.

Lister, J., Besing, J., and Koehnke, J. (2002). “Effects of age and frequency disparity on gap discrimination,: J. Acoust. Soc. Am., 111, 2793-2800.

Ludvigsen, C., C. Elberling, G. Keidser, T. Poulsen, (1990). “Prediction of intelligibility of non-linearly processed speech.” Acta Otolaryngol Suppl 469: 190-195.

Payton, K. L., L. D. Braida, S. Chen, P. Rosengard, R. Goldsworthy, (2002). Chapter 11. Computing the STI using speech as the probe stimulus. in Past Present and Future of the Speech Transmission Index. S. J. van Wijngaarden Ed. TNO Human Factors, The Netherlands: 125-138.

Ronnberg, J., Andersson, J., Samuelsson, S., Soderfeldt, B., Lyxell, B., and Risberg, J. (1999). “A speechreading expert: The case of MM,” J. Speech Hearing Lang. Res., 42, 5-20.

Nilsson, M., Soli, S.D., and Sullivan, J.A. (1994). "Development of the Hearing in Noise Test for the measurement of speech reception thresholds in quiet and in noise," J. Acoust. Soc. Am., 95, 1085- 1099.