Hearing Aid Research

RLE Groups
Sensory Communication Group

Sponsor
National Institutes of Health Grant R01 DC00117

Academic and Research Staff:
Professor Louis D. Braida, Dr. Joseph Desloge, Dr. Julie Greenberg, Dr. Andrew Oxenham, Dr. Karen L. Payton, Dr. Charlotte M. Reed

Visiting Scientists and Research Affiliates
Dr. Paul Duchnowski, Dr. Raymond Goldsworthy, Dr. Kenneth W. Grant, Dr. Christine Koh, Dr. Jean C. Krause, Dr. Christine M. Rankovic, Dr. Peninah S. Rosengard

Graduate Students
Shaoyan Chen, Andrew Dykstra, David Messing

Undergraduate Students
Jason Chiu, Michael Naber, Daniel L. Steele

Technical and Support Staff
Lorraine Delhorne, Joseph A. Frisbie, Denise Rossetti

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

Models of Speech Intelligibility

Analysis and development of intelligibility metrics

We have devoted substantial efforts towards developing speech-based Speech Transmission Index (sSTI, Houtgast and Steeneken, 1985) metrics applicable to a variety of processing conditions and listener populations. We have demonstrated that most previously-proposed sSTI metrics successfully predict intelligibility of speech corrupted by additive noise and reverberation (Payton and Braida, 1999; Goldsworthy and Greenberg, 2004), but that all of these metrics fail when applied to speech processed by non-linear operations (Payton et al, 2002; Goldsworthy and Greenberg, 2004; Goldsworthy, 2005) such as amplitude compression hearing aids, conventional noise reduction algorithms, and coding strategies used in CI speech processors. This is consistent with results from other researchers (Van Buuren et al., 1999; Ludvigsen et al., 1993; Drullman, 1995), who have also demonstrated the inadequacy of sSTI metrics to predict the intelligibility of speech processed by non-linear operations.

Our recent work (Goldsworthy and Greenberg, 2004; Payton et al., 2002) presented the first detailed analyses of *why* these sSTI metrics fail in the context of non-linear operations. By examining intermediate variables, we established explicit mathematical relationships among the various metrics, developed a comprehensive framework for comparing the different sSTI metrics proposed in the literature, and identified the issues that make all of these existing metrics unsuitable for non-linear operations.

This analysis also led to ideas for how to modify previously-proposed sSTI metrics for application to non-linear operations. In Goldsworthy and Greenberg (2004), we analyzed eight different sSTI-based intelligibility metrics, four previously proposed and four novel. The previously proposed metrics were based on: magnitude (Payton et al., 2002), phase-locking (Drullman et al., 1994b), envelope regression (Ludvigsen et al., 1990), and normalized covariance (Hollube and Kollmeier, 1996). Three of the four novel metrics were simple modifications of previously proposed metrics: modified magnitude, modified phase-lock, and modified envelope regression (MERM); the fourth was a preliminary version of the novel NCM (see paragraph at end of current section).

First, to ensure that they behave reasonably under the conditions for which the STI was originally intended, all eight metrics were compared to each other and to the traditional STI for conventional acoustic degradations (additive noise and reverberation). Seven metrics exhibited one-to-one correspondence with, and can be mapped to, the traditional STI. Only the normalized covariance metric failed to adequately predict the effect of additive noise and reverberation, so it was rejected and subsequent work only considered the other three previously-proposed metrics. Next, intermediate variables associated with all eight metrics were examined for two non-linear operations, spectral subtraction and envelope thresholding. Envelope thresholding was used because it is related to the N-of-M processing used in some CI speech processors. The intermediate variables associated with the four novel metrics all produced qualitatively reasonable results for the non-linear operations. This is in contrast to the three previously-proposed metrics whose intermediate variables often took on invalid values (outside the range 0-1) when applied to these non-linear operations. These results indicate that the four novel metrics produce qualitatively reasonable results and have potential for extending sSTI to non-linear operations while retaining applicability to conventional acoustic degradations.

The work in Goldsworthy and Greenberg (2004) inspired the development of the novel NCM that will be the focus of our proposed work. The NCM represents a substantial departure from previous sSTI metrics. It is relatively simple to calculate and eliminates many of the intermediate quantities (e.g., modulation metric, apparent SNR calculation) associated with traditional STI and sSTI metrics (Goldsworthy, 2005; Greenberg and Goldsworthy, 2005).

Application of sSTI to cochlear implants and noise reduction

Simulations of CI speech processing. We have implemented, in MATLAB (MathWorks, Inc.), a channel-vocoder simulation of CI speech processing. Many researchers have used similar systems to study CI processing effects with normal-hearing listeners (Dorman et al., 1997; Fu et al., 1998; Qin and Oxenham, 2003). Our simulation permits selection of a number of CI-processing parameters: the number, width, and corner frequencies of the bandpass filters, the method of envelope calculation (halfwave, fullwave, squaring, and Hilbert transform), and the lowpass cutoff frequency filter used to smooth the envelopes. For presentation to normal-hearing listeners, a broadband signal is reconstructed by summing together signals for each band that are determined from the corresponding envelope, a stationary noise carrier, and the same bandpass filters used in the analysis stage.

Some CI speech processing strategies are based on an approach known as N-of-M processing, where N out of M electrodes are active during each stimulation cycle (Loizou, 1998). For example, the SPEAK speech processing strategy (Cochlear Corp.) uses a proprietary algorithm to select N channels for stimulation in each processing frame, where N varies from frame to frame. The algorithm selects the N highest-energy bands in each frame; N varies from 5 to 10 and depends on the spectral composition of the input signal, with an average value of N=6 (Loizou, 1998). We simulate N-of-M processing by selecting, during each 4-ms frame, the N frequency bands with the highest energy and setting the remaining M-N envelopes to zero. In our simulation, N is a constant and an experimental variable, in other words, N is fixed for a particular condition.

Application of sSTI to hearing aids and hearing impairments.

We have developed two simulators for use in studying the effect of different hearing-aid processing schemes with hearing-impaired listeners. The first hearing aid simulator, implemented in MATLAB (MathWorks, Inc.), uses the dual front-end AGC (Moore and Glasberg, 1988), which includes a slow-acting wideband automatic volume control to determine the gain for most acoustic conditions, plus a fast-acting AGC to provide transient suppression. Our simulation of the dual front-end AGC includes two optional features, a hold-timer (Stone et al., 1999) and a signal-to-noise ratio (SNR) detector (Martin, 1992). The hold timer reduces pumping without extremely long recovery times, preventing gain fluctuations during speech and brief pauses in speech. The SNR detector modifies the release time of the slow acting AGC component so that it releases more quickly when strong speech (from the hearing-aid wearer's voice) is followed by weaker speech (from another talker). We used this simulator to implement five different algorithms, four dual front-end AGC algorithms (all combinations of with/without the hold timer and with /without SNR detection) plus a linear reference condition with compression limiting. We studied twenty acoustic conditions representative of everyday listening situations. Intelligibility and quality ratings for five hearing-impaired subjects showed no clear differences among the four dual front-end AGC algorithms. Major differences existed between the linear reference condition and the dual front-end systems; however, the direction of these differences varied with subject (Aguayo, 2001).

The second hearing aid simulator is implemented in Simulink (MathWorks, Inc.). It allows compression to be applied independently in up to four bands with variable static compression ratios, attack times, and release times. The simulator also has an option to include NAL-R or other frequency gain characteristics. The parameters of this simulation can be modified easily to mimic a wide range of hearing aids. This hearing aid simulator was used, along with an expansion hearing loss simulation (Duchnowski and Zurek, 1995; Lum and Braida, 2000), to study monaural linear-gain and WDR-MB compression hearing aids (Rosengard et al., 2005). The purpose of this study was two-fold: 1) to determine the extent to which four-channel WDR-MB compression can counteract the perceptual effects of reduced auditory dynamic range, and 2) to examine the relationship between objective measures of speech intelligibility and categorical ratings of speech quality for speech processed with amplitude compression. Intelligibility scores as well as subjective ratings of speech quality were collected for simulations of a flat, 50 dB, hearing loss and a sloping hearing loss. Results showed that moderate compression provides a small but significant improvement in speech intelligibility, relative to linear amplification, for simulated flat-loss listeners. This benefit was found for speech at conversational levels, both in quiet and with babble as additive noise. Sloping simulated-loss listeners did not demonstrate improvement in intelligibility with compression. However, in all listeners, ratings of pleasantness decreased as compression ratio increased.

Evaluating sSTI metrics with simulated hearing aids and hearing impairment. We evaluated three previously-proposed sSTI metrics for their ability to predict speech reception scores of normal-hearing subjects listening to speech processed by hearing-loss and WDR-MB compression hearing-aid simulations. The speech materials were nonsense sentences in quiet, in speech-shaped noise (SNR=5dB), or in restaurant babble (SNR=5dB). Two different hearing losses were simulated, a 50-dB flat loss and a moderate sloping loss. For the flat-loss, we examined a linear gain condition and two amplitude compression conditions, with compression ratios of 2 and 3. For the sloping loss we examined a linear gain condition and one amplitude compression condition, with compression ratios of 1.5 in the low frequency bands (<1.5 kHz) and 2.5 in the high frequency bands (>1.5 kHz). Attack time and release time constants for each band were 5 ms and 200 ms, respectively.

Preliminary attempts to calculate the previously-proposed sSTI metrics revealed that their intermediate variables often took on invalid values (greater than one), erroneously indicating that the hearing loss simulation increased intelligibility. This problem is consistent with limitations of existing sSTI metrics observed by others (e.g., Drullman, 1995) as well as our own theoretical analyses described above: the nonlinear expansion operation used in the hearing-loss simulation confounds the previously-proposed metrics. To address this problem, the hearing loss was modeled using a method previously used successfully in STI calculations, specifically, modeling the elevated thresholds as an internal noise that is added to the output of the hearing-aid simulation. This method has been shown to provide accurate predictions of intelligibility scores of actual hearing-impaired subjects (Payton et al., 1994; Humes et al., 1986). Accordingly, the sSTI results reported below all used internal noise to model the hearing impairment. Note that intelligibility was measured with the hearing-loss simulation using nonlinear expansion; only the sSTI calculation was impacted by this threshold modeling technique.

As a step in generating sSTI values, we first computed speech-based modulation transfer functions (sMTFs) for the hearing-aid and hearing-loss simulation conditions (Payton et al., 2002). For both the flat and sloping simulated losses with linear amplification, we were able to compute theoretical MTF values for comparison. The magnitude and phase-locked sMTFs matched the theoretical MTFs across all modulation frequencies and for all noise conditions, while the sMTFs generated by the Steeneken and Houtgast metric demonstrated deviations at the higher modulation frequencies (documented as a systemic problem with this technique in Payton et al., 1999). This metric also exhibited sMTF deviations at high modulation frequencies for all of the amplitude compression conditions. For the flat-loss amplitude compression conditions, the magnitude and phase-locked metrics produced sMTFs that were comparable and relatively flat as a function of modulation frequency. For sloping-loss conditions with the frequency-dependent amplitude compression, sMTFs computed for the magnitude metric were the only ones that produced valid values over the entire modulation frequency range considered. The Steeneken and Houtgast metric generated sMTFs greater than one for some modulation frequencies and the phase-locked sMTF produced negative values at certain modulation frequencies when the response envelope was out of phase with the input envelope.

To assess the predictions of the three previously-proposed sSTI metrics, we plotted intelligibility scores vs. sSTI values for measurements collected with normal-hearing subjects listening to speech processed by the hearing-loss and hearing-aid simulations. (When sMTF values outside the range 0-1 were produced, the values were clipped). Qualitatively similar results were observed for all three metrics. None of the metrics tested predicted STI values consistent with intelligibility scores across all processing conditions. All metrics predicted that linear processing conditions would be more intelligible than compression conditions, but the measured scores were similar across processing conditions for a fixed noise type. Within a class of processing (linear amplification or amplitude compression), we observed reasonable consistency between STI predictions and intelligibility (Payton et al. 2002).

Given that amplitude compression had a much greater impact on the sSTI metrics than on intelligibility, we investigated the idea of compensating the sSTIs to remove the affect of amplitude compression distortion from the metrics. For a given signal-to-noise ratio and compression ratio, we were able to determine the extent to which the sSTI metrics were reduced solely due to amplitude-compression distorting the intensity for the acoustic degradations and processing conditions of Payton et al. (2002). When the three previously-proposed sSTI metrics were normalized by this effective modulation reduction (EMR), the revised sSTI metrics calculated from stimuli with and without compression were almost identical and fit the intelligibility scores with a correlation coefficient of 0.88, compared to the initial correlation coefficient of 0.40 (Chen, 2005; Chen and Payton, 2005).

Implementing noise reduction techniques. We have extensive experience developing and evaluating signal processing algorithms for hearing aids and cochlear implants. Aspects of this work that are not directly applicable to the current grant application, e.g., feedback reduction (Maxwell and Zurek, 1995; Greenberg et al., 2000) and microphone-array hearing aids (Welker et al., 1997; Greenberg and Zurek, 2001; Greenberg et al., 2003; Koul and Greenberg, 2005), are not summarized here. Below we describe our previous work on signal processing algorithms for cochlear implants. These algorithms are intended as front-end noise reduction systems that process the microphone signals before they are applied to the input of the CI speech processor.

Spectral Subtraction: Spectral subtraction is a single-microphone noise reduction method based on subtracting an estimate of the noise spectrum from the spectra of a noisy input speech signal. We have implemented a generalized spectral subtraction algorithm (Lim and Oppenheim, 1979) based on the equation $P(F,i) = D(F,i) - \kappa N(F)$, where F is a frequency index, i is the frame index, $D(F,i)$ is the short-time magnitude spectrum of the input, $N(F)$ is the spectral estimate of the noise, $P(F,i)$ is the short-time magnitude spectrum of the output, and κ is a control parameter. $P(F,i)$ is recombined with the original phase and the time-domain output signal is reconstructed by performing overlap-add with 25-ms Hamming windows.

We have measured intelligibility in CI users as well as normal hearing subjects listening to 8- and 20-channel simulations of CI speech processing. Compared to a reference condition, spectral subtraction improved the intelligibility of CI processed speech by an average of 30 percentage points (Goldsworthy, 2005), confirming results from other work applying spectral subtraction to CI-processed speech (Weiss, 1993; Hochberg et al., 1992). This is in contrast to the established result that spectral subtraction does not improve speech reception for normal-hearing listeners (Lim and Oppenheim, 1979) and illustrates the need to tailor any robust intelligibility metric for different listener types. Section 3-c-iii describes the experimental conditions and the ability of sSTI to predict this improvement for CI-processed speech.

Binaural gain control for noise reduction: Noise reduction algorithms using more than one microphone are promising because they are able to exploit direction-of-arrival differences between signal and noise. In this work we developed and evaluated an algorithm based on binaural gain control. A number of researchers (Kollmeier and Koch, 1994; Lindemann and Melanson, 1997; Margo et al., 1997; Roman et al., 2003; Wittkop et al., 1997) have considered algorithms of this type where signals from two microphones, worn at the ears, are first added and then subjected to time- and frequency-dependent gain control. The time-varying gain for each frequency band is determined from characteristics of the two microphone signals; while the instantaneous SNR within a particular frequency band is unchanged, the algorithm suppresses bands with high noise levels in each time frame. This potentially improves intelligibility by suppressing noise components that would otherwise mask the desired signal at other frequencies.

Our binaural gain control algorithm follows the basic structure described by Kollmeier and Koch (1994) and Lindemann and Melanson (1997); that is, for each time frame, the gain for each frequency band is a function of the intermicrophone phase difference and the intermicrophone magnitude difference. Assuming that the listener is facing the desired speaker, frequency bands dominated by sources in front of the listener will have small intermicrophone phase and amplitude differences and receive no attenuation, while frequency bands dominated by noise from other directions will have larger intermicrophone differences and therefore tend to be attenuated. As in human hearing, intermicrophone phase differences are used at lower frequencies, while intermicrophone amplitude differences are also used at higher frequencies.

In comparison to previously described algorithms, our implementation of the binaural gain control algorithm uses higher frequency resolution and selects the gains based on a function that is more strongly dependent on intermicrophone amplitude and phase differences. The algorithm uses 32-ms frames (31 Hz frequency resolution) with 50% overlap. Below 800 Hz, the intermicrophone phase difference for each frequency bin is used to determine the gain applied at that frequency. Above 800 Hz, both intermicrophone amplitude and phase differences are used to determine gain value applied at each frequency.

We have performed two preliminary evaluations of the binaural gain control algorithm. In the first, speech plus babble was presented to three normal-hearing subjects and one CI user (Goldsworthy and Greenberg, 2000). This study did not include a simulation of CI speech processing. An adaptive method was used to determine the speech reception threshold (SRT), the SNR at which the subject identified 50% of the keywords. When compared to a reference condition consisting of the sum of the two microphone signals, the binaural gain control algorithm improved SRTs by 16 dB for the CI user and by 12-14 dB for the normal-hearing subjects. The second study (Goldsworthy, 2005; Goldsworthy and Greenberg, 2005) involved CI users as well as normal-hearing subjects listening to 8- and 20-channel simulations of CI speech processing. Compared to a reference condition, the binaural gain control algorithm improved the intelligibility of CI processed speech by an average of 35 percentage points, corresponding to an improvement in SNR of approximately 10 dB.

Evaluating sSTI metrics with CI simulations and noise reduction. We have performed a comprehensive evaluation of the intelligibility metrics described in above by comparing their predictions to speech reception scores for CI users as well as normal-hearing subjects listening to a simulation of CI speech processing under a variety of conditions. The SNRs listed below were used to test the normal-hearing subjects; for each CI subject, we adjusted the SNR by an offset based on preliminary measures of each subject's speech reception performance (Goldsworthy, 2005). Three CI users and six to eight normal-hearing listeners were used to evaluate the intelligibility of CUNY sentences (Boothroyd et al., 1985) under each of the following processing conditions:

Acoustic degradations: The target speech material was degraded by additive noise and/or reverberation. Reverberant impulse responses were generated by the image method (Allen and Berkeley, 1979; Peterson, 1986). Anechoic conditions included quiet plus three noise types (speech-shaped noise, multi-talker babble, or single talker time-reversed speech) at three SNRs (-3, 0, or 3 dB). Reverberant conditions included two levels of reverberation (T_{60} = 0.15 and 1.2 s) in quiet, in speech-shaped noise, or in time-reversed speech (SNR=3 dB). The degraded speech was presented through an 8-channel CI simulation for the normal-hearing subjects. All CI users all had Clarion processors.

N-of-M processing: The target speech was presented in quiet or in one of the three types of additive noise mentioned above at 0 dB SNR. The total number of channels in the simulation of CI speech processing was M=20, and the number of active channels was an experimental variable (N=3, 6, 9, or 20). CI users did not participate in this study.

Spectral subtraction: The target speech was degraded by additive speech-shaped noise and then processed by the spectral subtraction noise reduction algorithm. Conditions included eight values for the control parameter (κ = 0, 0.5, 1, 1.3, 1.6. 2, 4, 8), where κ = 0 is no processing and κ=1 is standard spectral subtraction. Normal-hearing subjects used two CI simulations: the SNR was 0 dB for the 8-channel CI simulation and -3 dB for the 20-channel CI simulation. CI users included subjects with Clarion and Nucleus processors.

Binaural noise reduction (two microphones): The target speech, located in front of the listener, was corrupted by additive noise (speech-shaped noise or single-talker time-reversed speech) located at 60 degrees azimuth in the same horizontal plane as the target speech and microphones. Acoustic conditions included anechoic and mildly reverberant (T_{60} = 0.15 s) environments. The binaural noise reduction algorithm was compared to a reference condition consisting of the sum of the two microphone signals. Normal-hearing subjects used two CI simulations: 8-channel (SNR = -3 dB) and 20-channel simulation (SNR = -6 dB). CI users included subjects with Clarion and Nucleus processors.

For all of the conditions described above, we computed the predicted intelligibility for seven of the eight metrics described above. (As mentioned previously, the normalized covariance metric was rejected.) A correlation analysis revealed that the seven metrics form three groups: The three previously-proposed metrics are similar to each other over the full set of conditions (r>=0.97). Three of the four novel metrics, including the MERM, are similar to each other (r>=0.98). Only the NCM produces substantially different predictions (Goldsworthy, 2005).

All metrics produced qualitatively reasonable results for the acoustic degradations. As expected, the three previously-proposed metrics failed to produce valid values for N-of-M processing and spectral subtraction, but surprisingly did produce valid values for binaural noise reduction. The four novel metrics (NCM, MERM, and two other modified metrics) produced valid values for all conditions. Comparison of predicted and measured intelligibility scores showed that for most conditions, NCM and the modified metrics correctly predicted major trends due to factors such as SNR, level of reverberation, and number of active channels in N-of-M processing. For spectral subtraction, the novel metrics successfully predicted trends of increased intelligibility for moderate values of the parameter κ and degraded intelligibility for large values of κ illustrating the potential utility of these metrics as a tool for refining noise reduction algorithms. However, of the four modified metrics only the NCM produced accurate predictions for the binaural noise reduction conditions.

The NCM exhibits great promise for predicting the intelligibility of degraded and/or processed speech. The NCM produced accurate predictions of speech reception for both CI users and normal hearing subjects listening to a simulation of CI processed speech for all the linear and nonlinear processing conditions tested. Using the mean-square-error and the correlation coefficient between observed and predicted scores as indicators of metric performance revealed that the NCM was exceptionally accurate in predicting speech reception performance for acoustically degraded speech after processing with spectral subtraction (mean-square-error: 5.04 rationalized arcsine units (RAU), correlation coefficient: 0.99) and binaural noise reduction (mean-square-error: 5.25 RAU, correlation coefficient: 0.97).

One limitation of the NCM and other novel metrics is that they generally do not predict the trends associated with noise type. For CI-processed speech, conditions with speech-shaped noise are more intelligible than modulated noise at equal SNRs (Qin and Oxenham, 2003; Goldsworthy, 2005), while for unprocessed speech, conditions with modulated noise are typically more intelligible (Festen et al., 1990; Peters et al., 1998; Nelson et al., 2003; Qin and Oxenham, 2003). The novel metrics exhibit no clear trends with respect to noise type, and thereby fail to predict observed trends for unprocessed or CI-processed speech. One goal of our proposed work is to develop a metric that can differentiate appropriately among noise types with the same long-term spectra but different short-time modulation characteristics.

Signal Processing for Hearing Aids.

Pilot Work on Vocoder-Based Frequency Lowering. In the first stage of this pilot research, we completed work on (a) the development of a software-based implementation of a vocoder-based frequency-lowering algorithm (Posen et al., 1993) which offers flexibility in the selection of analysis and synthesis bands in terms of their center frequencies and bandwidths and which can be tailored to individual hearing loss characteristics; (b) the application of this signal-processing scheme to a set of 768 C_1VC_2 syllables with three vowels V=/i a u/ and 16 initial consonants C_1=/p t k b d g f th s sh ch v tx z zh j/; and (c) the development of software for conducting identification experiments and data analysis.

In the second stage of the pilot research, we collected a set of data based on extensive, in-depth testing (totaling roughly 50 hours) of an individual with high-frequency hearing loss (data collection was begun on a second listener whose time constraints did not him permit to complete the test protocol). The hearing loss in the listener's test ear (the right ear) progressed from 30 dB HL at 1 kHz to 75 dB HL at 4 kHz (although no evidence was found for "dead regions" in the area of maximum loss). Her ability to identify initial consonants in nonsense syllables was examined for a set of baseline conditions and for two frequency-lowering conditions. The same protocol was followed for all conditions: a pre-training test score was obtained without correct-answer feedback; training was conducted on the identification task using trial-by-trial correct-answer feedback; and once asymptotic performance was achieved on training, post-training testing was conducting without feedback. (Separate sets of speech tokens were employed for training with feedback and for testing without feedback.)
To establish a baseline for the listener's consonant-identification ability, three conditions were examined: low-pass filtering at 1 kHz (LPF 1K), low-pass filtering at 2 kHz (LPF 2K), and linear amplification applied using Moore and Glasberg's (1998) Cambridge prescription (LA_Camb). Two frequency-lowering conditions were tested employing four analysis bands located in the region of 1.5-5.6 kHz, each of which controlled the output of a corresponding synthesis band located in the region of 0.5- 2 kHz (selected to lie along the slope of the hearing loss and into the region of good residual hearing). In the first frequency-lowering condition, frequency lowering was applied to the unprocessed speech signal (FL+UP) and in the second condition it was applied to linearly amplified speech (FL+LA_Camb).

The results of the post-training identification tests conducted with each of these five conditions were summarized in terms of overall %-correct, %-overall information transfer (IT), and %-feature IT for features of voicing, manner, and place of production. Performance was similar for the LPF 1K and LPF 2K conditions for overall percent correct (roughly 52%-correct) and for each of the other measures, indicating that no information was gained in the region of 1-2 kHz. Performance improved substantially, however, for linear amplification in the high-frequency region (LA_Camb) where overall performance was 81%-correct and feature scores ranged from 69%-IT for place to 72%-IT for manner to 89%-IT for voicing. Frequency lowering applied to the unprocessed speech signal (FL+UP) yielded overall %-correct performance similar to that obtained under the low-pass-filtering conditions (roughly 54%); performance on each of the three features, however, was somewhat improved with frequency lowering. Frequency lowering applied to speech with the Cambridge formula yielded overall %-correct performance (76%) that was equivalent to that obtained with the Cambridge formula alone (a t-test yielded no significant difference between the two conditions); however, this condition (FL+LA Camb) led to an improvement in voicing performance and a decrement in place performance compared to LA_Camb alone. Despite the magnitude of her hearing loss at high frequencies, this listener derives benefit from high-frequency amplification (unlike results reported by Hogan and Turner, 1998 for certain listeners with similar degrees of loss who show no improvements with high-frequency amplification).

The pilot results obtained here are promising for future research. The combination of vocoder-based frequency lowering with high-frequency amplification is a novel system that deserves to be explored with a wider variety of high-frequency hearing impairments, particularly those listeners who do not demonstrate benefit from high-frequency amplification (which may perhaps be associated with the existence of high-frequency “dead regions”). Even in the listener tested here (who did benefit from conventional high-frequency amplification), improvements were noted with the addition of frequency lowering in the area of voicing reception. Furthermore, the system is highly flexible in that analysis and synthesis bands can be selected to maximize the performance of individual listeners.

References

Allen, J.B. and Berkley, D.A. (1979). "Image method for efficiently simulating small-room acoustics", J. Acoust. Soc. Am. 65, 943-950.

Dorman, M.F., Loizou, P.C., Rainey, D. (1997). "Speech intelligibility as a function of the number of channels of stimulation for signal processors using sine-wave and noise-band outputs", J. Acoust. Soc. Am. 102, 2403-2411.

Drullman, R. (1995). "Temporal envelope and fine structure cues for speech intelligibility", J. Acoust. Soc. Am. 97, 585-592.

Drullman, R., Festen, J., and Plomp, R. (1994). "Effect of reducing slow temporal modulations on speech reception," J. Acoust. Soc. Am., 95, 2670-2680.

Festen, J.M., Van Dijkhuizen, J.N., and Plomp, R. (1990). "Considerations on adaptive gain control in hearing aids", Acta Oto. Suppl. 469, 196-201.

Fu, Q.-J., and Nogaki, G. (2004). “Noise Susceptibility of Cochlear Implant Users: The Role of Spectral Resolution and Smearing,” .J.A.R.O. 6, 19-27.

Goldsworthy, R.L. and Greenberg, J.E. (2000). "Algorithms used for noise reduction", Am. Speech-Language-Hearing Assoc. Audiology Conf., San Francisco.

Goldsworthy, R.G. and Greenberg, J.E. (2003). "Using the speech transmission index to predict the intelligibility of cochlear implant speech processing", Conference on Implantable Auditory Prostheses, Pacific Grove, CA, August 2003.

Goldsworthy, R.L. and Greenberg, J.E. (2004). "Analysis of speech-based Speech Transmission Index methods with implications for non-linear operations", J. Acoust. Soc. Am. in press.

Goldsworthy, R.L. (2005). "Improving Speech Reception in Noise for Cochlear Implant Users", Ph.D. Thesis, Harvard-MIT Division of Health Sciences and Technology, Massachusetts Institute of Technology, Cambridge, MA.

Hochberg, I. Boothroyd, A., Weiss, M. and Hellman, S. (1992). "Effects of noise and noise suppression on speech perception by cochlear implant users", Ear and Hearing 13(4), 263-271.

Houtgast, T. and Steeneken, H.J.M. (1985). "A review of the MTF concept in room acoustics and its use for estimating speech intelligibility in auditoria," J. Acoust. Soc. Am., 77, 1069-1077.

Humes, L.E., Dirks, D.D., Bell, T.S., Ahlstrom, C. and Kincaid, G.E. (1986). "Application of the articulation index and the speech transmission index to the recognition of speech by normal-hearing and hearing-impaired listeners", J. Speech Hearing Res. 29, 447-462.

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

Kollmeier, B. and Koch, R. (1994). "Speech enhancement based on physiological and psychoacoustical models of modulation perception and binaural interaction", J. Acoust. Soc. Am. 95, 1593-1602.

Krause, J.C., and Braida, L.D. (2004a). "Acoustic Properties of Naturally Produced Clear Speech at Normal Speaking Rates," J. Acoust. Soc. Am. 115, (1), pp. 362-378

Lim J.S. and Oppenheim, A.V. (1979). "Enhancement and bandwidth compression of noisy speech", Proc IEEE 67, 1586-1604.

Lindemann, E. and Melanson, J.L. (1997). "Noise reduction system for binaural hearing aid", U.S. Patent No. 5,651,071.

Loizou, P.C. (1998). "Mimicking the human ear", IEEE Signal Processing Magazine 15, 101-130.

Loizou, P.C., Dorman, M., and Tu, Z. (1999). "On the number of channels needed to understand speech," J. Acoust. Soc. Am. 106 (4), 2097- 2013.

Ludvigsen, C. (1987). "Prediction of speech intelligibility for normal-hearing and cochlearly hearing-impaired listeners." J. Acoust. Soc. Am., 82(4), 1162-71.

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

Margo, V., Schweitzer, C., and Feinman, G. (1997). "Comparisons of Spectra 22 performance in noise with and without an additional noise reduction preprocessor", Seminars in Hearing 18(4), 405-415.

Martin, G.W. (1992). "Real-time multiband adaptive hearing aids," S.M. Thesis MIT.

Moore, B.C.J. and Glasberg, B.R. (1988). "A comparison of four methods of implementing automatic gain control in hearing aids," Brit. J. Audiol., 22, 93-104.

Nelson, P.B., Jin, S.H., Carney, A.E., and Nelson, D.A. (2003). "Understanding speech in modulated interference: Cochlear implant users and normal hearing listeners," J. Acoust. Soc. Am. 113, 961-968.

Nelson, P.B. and Jin, S.-H. (2004). " Factors affecting speech understanding in gated interference: Cochlear implant users and normal-hearing listeners." J. Acoust. Soc. Am. 115 (5), 2286-2294.

Payton, K. L. and L. D. Braida (1999). "A method to determine the speech transmission index from speech waveforms," J. Acoust. Soc. Am., 106(6), 3637-48.

Payton, K.L., Braida, L.D., Chen, S., Rosengard, P., and Goldsworthy, R. (2002). "Computing the STI using speech as a probe stimulus," in T. Houtgast and H. Steeneken (eds) Past, present, and future of the Speech Transmission Index, TNO Human Factors, Soesterberg, The Netherlands, 125-138.

Payton, K.L., Uchanski, R.M., and Braida, L.D. (1994). "Intelligibility of conversational and clear speech in noise and reverberation for listeners with normal and impaired hearing," J. Acoust. Soc. Am., 95,1581-1592.

Stone, M.A., Moore, B.C.J., Alcantara, J.I., and Glasberg, B.R. (1999). "Comparison of different forms of compression using wearable digital hearing aids", J. Acoust. Soc. Am. 106, 3603-3619.

Weiss, M.R. (1993). "Effects of noise and noise reduction processing on the operation of the Nucleus-22 cochlear implant processor", J. Rehab. Res. 30(1), 117-128.

Wittkop, T., Albani, S., Hohmann, V., Peissig, J., Woods, W.S., and Kollmeier, B. (1997). "Speech processing for hearing aids: noise reduction motivated by models of binaural interaction", Acustica 83, 684-699.

PUBLICATIONS

Journal Articles

Koul, A. and Greenberg, J.E. (2005). "Using intermicrophone correlation to detect speech in spatially-separated noise," EURASIP Journal on Applied Signal Processing, special issue on Advances in Multimicrophone Speech Processing.

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.

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," J. Speech and Hearing Res. 48 (3) 702-714.

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

Meeting Papers Presented

Chen, S. and Payton, K. L. (2005). “Compensating the STI to predict intelligibility for amplitude-compressed speech”. in Proc. 2005 IEEE Workshop Appl. Sig. Process. Audio Acoust., New Paltz, NY, IEEE, 49-52.