Hearing Aid Research

Sponsor
National Institutes of Health Grants R01 DC00117, R01 DC007152.

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, Arlene Wint

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

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

This research is 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 listeners with real and simulated hearing loss are thus roughly equivalent in terms of their levels of audibility. 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.

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

The effects of audibility and age on masking for sentences in continuous and interrupted noise were examined in listeners with real and simulated hearing loss (Desloge et al, 2010). The absolute thresholds of each of ten listeners with sensorineural hearing loss were simulated in normal-hearing listeners through a combination of spectrally-shaped threshold noise and multi-band expansion for octave bands with center frequencies from 0.25-8 kHz. Each individual hearing loss was simulated in two groups of three normal-hearing listeners (an age-matched and a non-age-matched group). The speech-to-noise ratio (S/N) for 50%-correct identification of Hearing In Noise Test (HINT) sentences was measured in backgrounds of continuous and temporally-modulated (10 Hz square-wave) noise at two overall levels for unprocessed speech and for speech that was amplified with the NAL-RP prescription. The S/N in both continuous and interrupted noise of the hearing-impaired listeners was relatively well-simulated in both groups of normal-hearing listeners. The largest differences in performance between listeners with real and simulated hearing impairment occurred for speech under NAL-RP processing in the higher-level noise and was due primarily to the performance of two of the hearing-impaired listeners. For the most part, however, release from masking (the difference in S/N obtained in continuous versus interrupted noise) appears to be determined primarily by audibility. Age effects were not observed in this small sample. Observed values of masking release were compared to predictions derived from intelligibility curves generated using the Extended Speech Intelligibility Index (ESII; Rhebergen et al., 2006). The ESII-based predictions of masking release showed a tendency to over-predict observed values in the range of 0-5 dB and to under-predict for higher observed values of masking release (i.e., >7 dB).

1-B-2. Temporal modulation transfer functions.

Temporal modulation transfer functions (TMTFs) were obtained in nine listeners with moderate to profound sensorineural hearing loss (Desloge et al., 2009). The hearing loss of each impaired listener was simulated in a group of three age-matched normal-hearing listeners through a combination of spectrally-shaped masking noise and multi-band expansion. TMTFs in both groups of listeners were measured in a broadband noise carrier as a function of modulation rate in the range of 2 to 1024 Hz using a 3I, 2AFC procedure. The presentation level for the unmodulated broadband noise (whose duration was 500 ms) was set to be the maximum of either 70 dB SPL or the level such that noise was 10 dB above the lowest hearing threshold. The listeners with simulated hearing loss thus received signals at roughly the same level as their hearing-impaired counterparts. The shapes of the TMTF curves (defined as the measured threshold of modulation in dB as a function of frequency of modulation) were compared for listeners with real and simulated hearing loss. The simulated-loss subjects never did better than subjects with real hearing impairment or with no hearing loss. With the exception of one subject, the simulated-loss subjects behaved more closely to the HI subjects with more severe hearing loss. For HI subjects with less severe hearing loss, the AM-SIM groups actually perform worse than the corresponding HI subject. Several factors may be related to the inferior performance of the AM-SIM groups. These include (i) the level of the broadband noise stimulus itself, (ii) the shape of the hearing loss, and (iii) interactions between the stimulus noise and the simulation masking noise.

1-B-3. Forward-masking functions.

Forward-masking functions were obtained in seven listeners with moderate to severe sensorineural hearing loss (Braida et al., 2009). The hearing loss of each impaired listener was simulated in a group of three age-matched normal-hearing listeners through a combination of spectrally-shaped masking noise and multi-band expansion. Forward masking was measured in both groups of listeners for probe signals at frequencies of 500, 1000, 2000, and 4000 Hz using an on-frequency masker and two off-frequency maskers (0.55 and 1.15 times the signal frequency) under a 3I, 2AFC procedure. The probe signal was presented at 5 dB SL; signals and maskers were gated with 5-msec on/off times with a steady-state duration of 0 ms for probe signals and 100 ms for maskers; and values of masker-offset time to signal-onset time were in the range of 0 to 100 ms. For three of the seven HI listeners, the hearing-loss stimulation generally provided a good match to forward-masking results at a majority of the 12 probe frequency/masker frequency conditions. Slopes of the masking functions for both on-frequency and off-frequency maskers (and their corresponding slope ratios) were similar under real and simulated hearing loss, both in cases from which an active basilar-membrane compressive function may be inferred and in cases where such a function is inferred to be absent. For the remaining four subjects, greater variability was observed in the ability of the hearing-loss simulation to reproduce the results of a given HI listener. At some probe frequency/masker frequency conditions (which differed across HI subjects), good correspondence was obtained between the temporal-masking curves for real versus simulated hearing loss. At other conditions, however, the correspondence was poor and led to large differences in slopes and slope ratios for these conditions. Poor correspondence between temporal-masking functions for HI listeners and their corresponding AM-SIM groups was often due to early saturation of the forward-masking curves. In some cases, such saturation effects were observed in the data of the HI listeners and in other cases in that of the simulated–loss listeners.

1-B-4. Gap-detection.

Gap-detection thresholds were measured in listeners with real and simulated hearing loss under conditions of auditory or tactile presentation (Reed et al., 2009). The audiometric thresholds of each of ten listeners with sensorineural hearing loss (21-69 years of age) were simulated in groups of age-matched normal-hearing listeners through a combination of spectrally-shaped masking noise and multi-band expansion. The leading and trailing markers for the gap-detection task were 250- and 400-Hz sinusoids with a nominal duration of 100 ms. Gap-detection thresholds for a nominal baseline gap of 0 ms were measured for four different combinations of leading and trailing markers (250-250, 250-400, 400-250, and 400-400 Hz) using a 3I, 2AFC procedure. Auditory measurements were obtained for monaural presentation over headphones using a marker level set to be equal to the maximum of 70 dB SPL or 10 dB SL. Tactile measurements were obtained using sinusoids presented to the left middle finger through an Alpha-M AV-6 vibrator at a level of 25 dB SL. Results indicated that: (i) A strong correlation between auditory and tactile gap-detection thresholds was observed in only one condition (400-400 Hz) for the hearing-impaired listeners but not for the simulated-loss listeners. (ii) Spectral disparity between leading and trailing markers led to higher gap-detection thresholds in both the auditory and tactile modalities for listeners with both real and simulated hearing loss. The effects of spectral disparity in the auditory conditions were somewhat more pronounced for the simulated-loss listeners. (iii) Only modest correlations were observed between the gap-detection thresholds of individual HI listeners and their corresponding age-matched simulated-loss groups.

1-C. Significance

Our hearing-aid 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.

1-D. Plans for Upcoming Year

We will complete the analysis of data collected on 10 HI listeners and a total of 30 NH listeners with simulated hearing-loss and prepare manuscripts for publication on the following topics: temporal-modulation transfer functions, forward masking, and gap-duration detection. We will conduct an analysis correlating results obtained on the speech-reception study with those obtained on measures of psychoacoustic and cognitive abilities. We will complete work on manuscripts concerned with a review of the literature on the role of audibility in cochlear hearing loss in the areas of intensity perception, spectral resolution, and 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

Progress has been made towards integrating psychophysical measures using acoustic and electrical stimulation. One goal of the project is to develop improved methods of stimulation for cochlear implants. Cochlear implant recipients exhibit a wide range of auditory benefits that can be attributed to factors that differ across recipients such as etiology of deafness, age of implantation and early language development. While this performance range is generally acknowledged, there have been few attempts that specifically consider developing stimulation strategies that are tailored to individual recipients. A first step towards accomplishing such a tailoring would be an assessment of acoustic and electric psychophysics, as well as speech reception, in a group of recipients. Towards that end, we have commenced pilot studies on a battery of psychophysical measures using both acoustic and electric stimulation. The acoustic stimuli have been described in prior reports; we briefly describe electrical stimulation paradigms.

Pilot experiments have explored the psychophysics of temporal pitch perception. A typical stimulus used in these experiments is a 400ms pulse train that is controlled along a dimension of interest to probe a subject’s difference limens (DLs). In a simple frequency DL experiment, the target stimulus is generated as per the reference, but using a pulse rate that is higher in frequency. This example represents perhaps the simplest rate comparison that can be made in cochlear implant psychophysics and serves as a template as we discuss various dimensions or rate perception.

A variation on this measure considers the effect of stimulus intensity on the perception of pitch. Figure 1 illustrates a typical target and reference comparison for pulse trains of different amplitudes. A corresponding psychophysical measure would be determined as in the template measure, except the intensity of the target and reference stimuli are roved within a nominal range, for example between 60 and 100% of the subject's electrical dynamic range.

Another variation of the basic template considers the effect of modulation on the salience of a pitch cue associated with temporal fine structure. A straightforward modification of the rate pitch measure is to modulate both the reference and target. There are innumerable combinations of such, but we are considering the combination of using a 110Hz modulator on carrier frequencies of 440, 880, and progressively higher frequencies should subjects show an ability to complete the task. Figure 2 illustrates the basic comparison for exploring this dimension. (The frequencies of the stimulation and of the modulation have been lowered for ease of viewing.) This dimension is being explored in terms of the shape and rate of modulation in order to understand how these parameters affect the salience of any pitch cue that can be derived from the temporal fine structure.

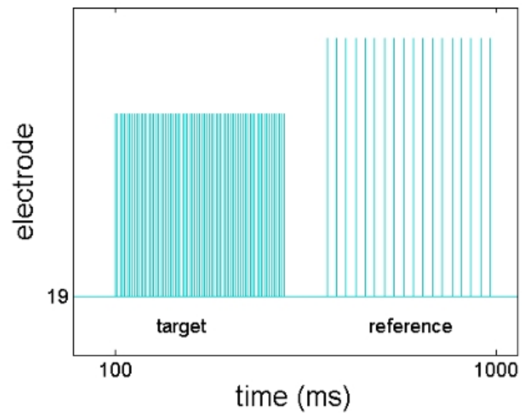

Figure 1.

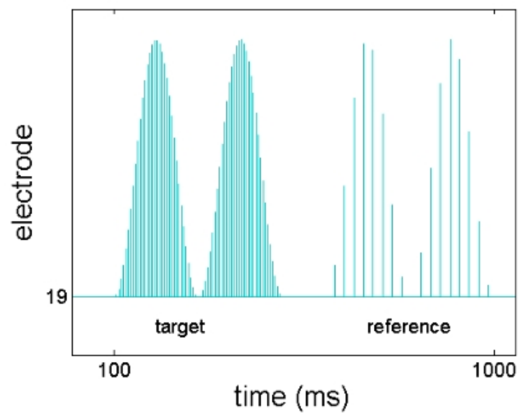

Figure 2.

Another dimension being explored in the perception of rate pitch is the effect of stochastic firing. Studies of neural synchrony show the ability of the auditory nerve to phase lock to sound as high as several thousand Hz. But this synchrony does not imply that every nerve fires to every cycle of the incoming sound. On the contrary, it has been shown that for high frequency sounds, a given neuron fires in synchrony to a phase of the sound, but not to every cycle. It is argued that higher level neural mechanisms probably take advantage of ensemble convergences to register a salient percept. The electrical stimulation provided by a cochlear implant might benefit by stimulating more stochastically for higher frequency sounds. For example, at rates greater than 1000Hz, the strategy might purposefully drop a percentage of the encoded pulses. To explore the potential benefits or drawbacks of doing such, we considered the basic rate pitch template under rules of dropping pulses. One rule for doing so would be to draw spikes from a Poisson distribution, but maintaining the fundamental rate of stimulation. For example, Figure 3 illustrates a target and reference comparison for spike times drawn from a Poisson distribution based on using a fundamental period of 1/F for the reference and 1/2F for the target. The resulting reference and target stimuli have a stochastic distribution of timings, but the average rate of the target is twice that of the reference. This approach yields insight in potential benefits of specifically allowing stochastic behavior for high frequencies in the temporal fine structure when the auditory nerve may not be able to completely synchronize to a constant rate pulse train. Although the Poisson distribution may not be the ideal distribution to use, as it produces a large temporal gaps which are perceived with a noise like quality, it serves as a starting point for exploring this dimension.

Another relevant dimension of exploring rate pitch in cochlear implants is to examine the effects of across electrode masking. This dimension is crucial to our understanding of how well temporal fine structure can be encoded in one electrode region, without disturbing such information encoded in another region. A basic measure for exploring this effect is to modify the rate pitch measure to include a masker on another electrode. For example, Figure 4 illustrates the presence of a pulse train on electrode 19 and a masker on electrode 12. If this was the reference stimulus, then the target would have an identical masker, but the pulse train on electrode 19 would have an adaptively higher rate. Again, for exploring this dimension there are innumerable combinations of maskers, references, and targets. Out initial measures focused on the use of a masker that has random firing with pulse periods drawn from rates between 400 and 1600Hz. We examined how the masking pattern changes with electrode position as the masker is moved closer to the electrode that the rate pitch comparison is specified. For example, an initial test explored using a rate pitch comparison on electrode 19 at rates of 110, 220, 440Hz and with maskers tested at electrode 18, 16, 13, and 9.

Pilot studies have explored some of the nuances of how temporal fine structure interacts with intensity, modulation, stochastic firing, and masking. The collected data is being analyzed and a larger study that considers the relationship between psychophysics of pitch perception using electrical and acoustical stimuli, and relations with speech reception measures, is being designed.

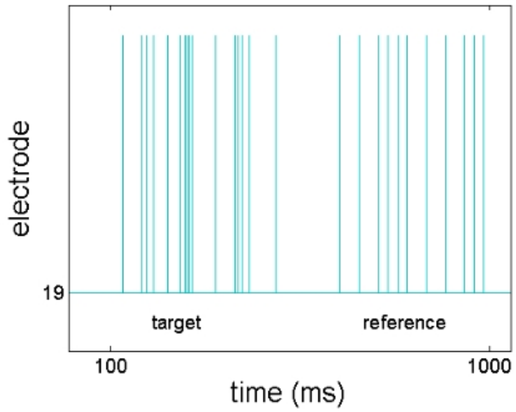

Figure 3.

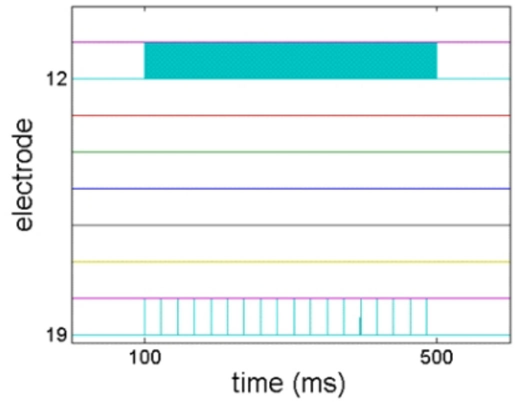

Figure 4.

The focus of this grant is to develop and assess models that accurately predict the effects of hearing impairment and cochlear implantation on speech comprehension as well as on simple psychoacoustic measures. The focus of the ARRA Supplement is to incorporate more basic measures of cochlear implant psychophysics into such models.

Towards this end, we have begun pilot experiments that investigate the perception of temporal fine structure. Cochlear implants can convey pitch in two distinct ways: 1) the place of stimulation and 2) the rate of stimulation. Conventional sound processing strategies make use of the first mechanism in that higher frequencies of sound are processed to stimulate more basally implanted electrodes, and lower frequencies sounds to more apically implanted electrodes. The notion of using the rate of stimulation to convey either a secondary, or nuanced, pitch percept has been entertained by others; however, a successful processing strategy has not yet been incorporated into commercial strategies.

Other research laboratories have shown that cochlear implant recipients can use electric stimulation rate as a pitch cue for stimulation rates less than 400 Hz. In contrast, physiological studies have shown that the auditory nerve in mammals can respond synchronously to electrical stimulation rates well above 2000 Hz. It is unknown why implant recipients' ability to use rate pitch rapidly deteriorates above 400 Hz given the physiological data suggesting that high rate synchrony is preserved at the level of the auditory nerve.

Our initial studies considered the role of auditory training on perceiving this cue. We reasoned that implant recipients have been deprived of such rate pitch cues for the duration of their deafness, so perception of rate pitch may require substantial training. As the pilot testing proceeded, however, it became clear that subjects could perceive a difference between different stimulation rates, even when both rates were above 1000 Hz. The basic test of this ability was a two-interval, two-alternative, forced-choice experiment. The standard stimulus was a 400 ms pulse train presented on a single electrode, and the target was identical except for having an adaptively controlled higher stimulation rate. The initial value of the rate difference was one octave. The procedure adapted, requiring the subject to answer correctly twice in a row before making the task more difficult. The procedure was designed to converge at a point in which the subjects could correctly identify the stimuli with 70.7% accuracy.

Figure 9 illustrates performance on this rate pitch measure for one subject. The rate difference limen is given in semitones (a semitone being 1/12th of an octave, a frequency ratio of 1.0595). At frequencies less than 400 Hz, the subject performs with a resolution better than two semitones. Consistent with the literature, this resolution deteriorates between 400 and 1000 Hz. However, the subject maintained an ability to perform the task, and with increasing accuracy over the ranges of 1000 to 4000 Hz.

This phenomenon has not been studied to any extent before. Part of the difficulty of investigating this phenomenon is that the increased stimulation rates require careful loudness balancing. Our current hypothesis is that this subject may have been using a correlated loudness percept to allow him to discriminate the stimuli correctly. We are currently modifying the experiment paradigms in an attempt to isolate the loudness and pitch percepts that are associated with the increase in stimulation rates.

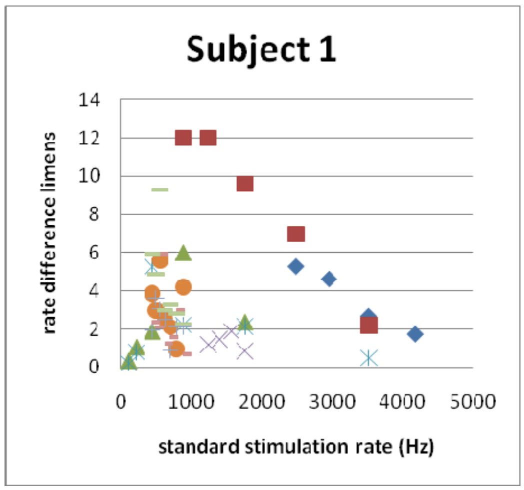

Figure 5.

II-C. STI Based Models of Speech Intelligibility

The primary effort was focused on improving the short-time STI STSTI) developed previously. Payton and Shrestha (2008a,b) presented STSTI algorithms based on the Envelope Regression and Normalized Correlation STI methods. The algorithms successfully track the intelligibility of speech samples in noisy and reverberant environments using rectangular windows as short as 1/3 s. This lower bound was based primarily on the fact that the various methods began to deviate from the Theoretical STI for windows shorter than this. Figure 6 shows the STSTI computed based on the Envelope Regression method and using 1/3 s and 80 ms windows. Ideally, there should be a 1:1 correspondence between the Envelope Regression Method and the Theoretical Method. However, the linear relationship between the two grows weaker as window length is decreased below 1/3 s as shown by the right panel of Fig. 6.

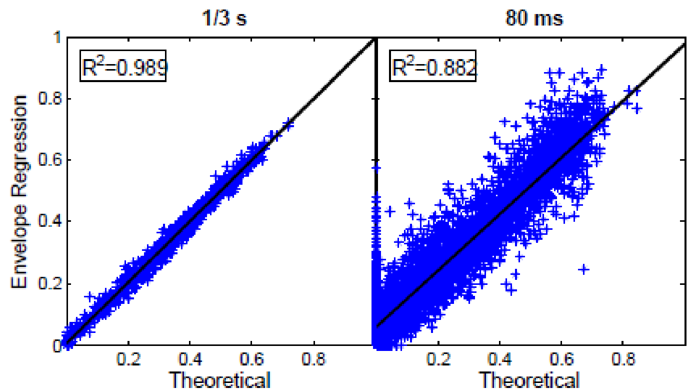

Figure 6. STSTI computed from the Envelope Regression Method vs. STI from the Theoretical Method for the condition of 0 dB SNR speech shaped noise. The left and right panels correspond to the standard STSTI computed using 1/3 s and 80 ms windows, respectively. The solid line indicates the line of best fit for the data in each panel.

The current study is aimed at understanding and improving the performance of the STSTI so it can be used to predict the intelligibility of speech for windows shorter than 1/3 s. Work to date has considered both Envelope Regression and Normalized Covariance methods but, since they exhibit similar behaviors across conditions, the results presented herein will be for the Envelope Regression Method in the condition of 0 dB SNR speech-shaped noise only. A better understanding in simple conditions like additive noise or reverberation could serve as a foundation for adapting the STSTI to non-linear and hearing aid processed speech. Reducing the window length, perhaps to the phoneme level (e.g. a few tens of milliseconds), could provide some insight into the way certain phonemes contribute to the intelligibility of speech. This could in turn be used as a tool to train talkers to produce more intelligible speech or as means to understand the differences in intelligibility between normal and hearing impaired listeners. In order to achieve these goals, the two primary issues that are being addressed are non-zero STI values produced during silence periods (i.e. the data points that fall along the vertical axis of the right panel in Fig. 6) and deviation from the Theoretical Method over the entire range of the STI.

During silence periods, non-zero intelligibility predictions from the STSTI methods are generated by a statistical similarity between the envelopes of clean and degraded speech. In its standard form, the STSTI is not able to determine whether or not a window consists of speech or silence. To resolve this issue, a speech/silence detection algorithm, referred to as the Adaptive Mean Algorithm (AMA), was developed. In theory, if a portion of the stimuli consists of silence, the intelligibility, resulting STI and modulation metrics from which the STI is derived should all be zero. The primary function of the AMA is to determine which windows contain silence and set the modulation metrics of those windows to zero. This decision is based on whether or not the ratio of the window mean to an approximation of the long term mean of the clean speech envelope falls below a threshold. As shown by Fig. 7, the AMA corrects nearly all of the non-zero STI values that occurred during silence periods.

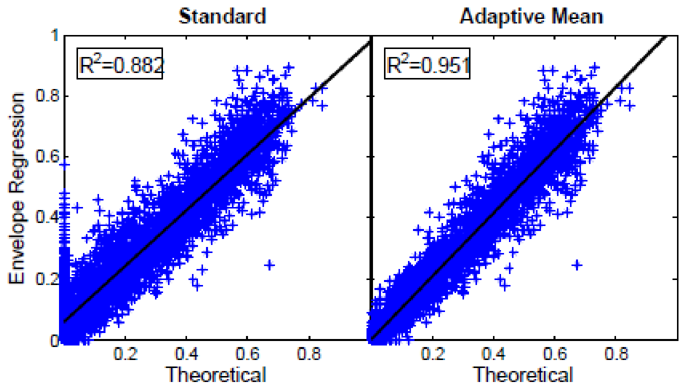

Figure 7. STSTI computed from the Envelope Regression Method vs. STI from the Theoretical Method for the condition of 0 dB SNR speech shaped noise using an 80 ms window. The left and right panels correspond to the STSTI computed using the standard approach and AMA, respectively. The solid line indicates the line of best fit for the data in each panel. The oval encircles silent periods when the standard STSTI generates non-zero results.

The current study also focused on determining the cause of the deviation from the Theoretical Method over the entire range of STI values. It was initially thought that this could have been the result of pre-processing the intensity envelopes (e.g aliasing brought on by downsampling). To test this hypothesis, the STSTI was computed without downsampling. Results showed that aliasing was negligible and downsampling had virtually no effect on performance. The next step was to examine the effect lowpass filtering during envelope extraction had on the behavior of the STSTI. An initial test was performed by computing the STSTI from the squared octave band speech signals by removing the lowpass filter from the envelope extraction stage. Preliminary results show that this greatly improves performance. Figure 8 illustrates this.

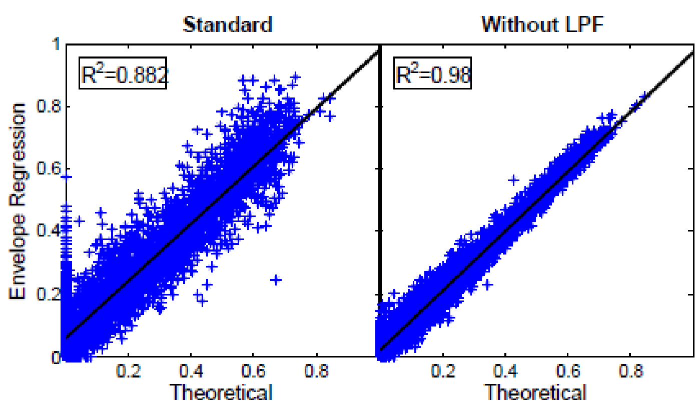

Figure 8. STSTI computed from the Envelope Regression Method vs. STI from the Theoretical Method for the condition of 0 dB SNR speech shaped noise using an 80 ms window. The left panel corresponds to the STSTI computed using the standard approach. The right panel corresponds to the STSTI computed without lowpass filtering. The solid line indicates the line of best fit for the data in each panel.

The cause of this is not fully understood, but the following hypothesis has been made. As window length is decreased, the segments of clean and degraded speech begin to resemble constants. This results in a significant decrease in the statistical values (e.g. variance, covariance, correlation etc.) of the envelopes, from which the modulation metrics are based. Consequently, the modulation metrics become unstable in the sense that they take on a wider range of values, particularly those outside the valid 0-1 range. The addition of temporal fine structure through the removal of lowpass filtering seems to reverse this effect by causing an increase in the statistical parameters which in turn results in a higher percentage of modulation metrics that fall in the valid range.

When the AMA and the addition of fine structure are used together, the STSTI gives the best overall performance. An example of this is shown in Fig. 9. In general, the combination of both modifications results in a reduced and more uniform deviation from the Theoretical Method for the entire range of STI values. The net effect is a closer linear relationship between the Theoretical and STSTI methods which corresponds to more accurate STI values.

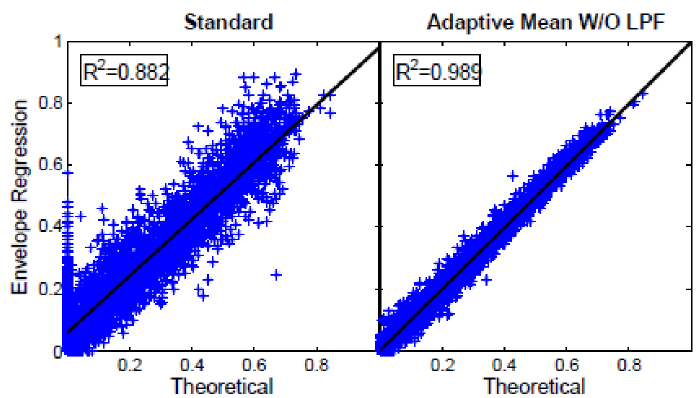

Figure 9. STSTI computed from the Envelope Regression Method vs. STI from the Theoretical Method for the condition of 0 dB SNR speech shaped noise using an 80 ms window. The left panel corresponds to the STSTI computed using the standard approach. The right panel corresponds to the STSTI computed using both modifications. The solid line indicates the line of best fit for the data in each panel.

These results demonstrate that the STSTI has the potential to be used as a reliable predictor of intelligibility for windows shorter than 1/3 s. However, more work is still needed to better understand why the addition of temporal fine structure improves STSTI performance. To do this, an analysis is being carried out that focuses on how the behavior of the intermediate STSTI metrics change as a function of window length and lowpass filter cutoff frequency. The modifications will also be studied in the presence of a fluctuating masker such as gated noise to examine the accuracy of the STSTI in conditions with non-stationary background noise. The validity of the modifications will also be examined by comparing the STI values with word scores from actual listener tests. Once these results are obtained, a more precise window length limitation for the modified STSTI can be determined.

Publications

Desloge, J. G., Reed, C. M., Braida. L. D., Delhorne, L. A. and Perez, Z. D. (2010)
"Temporal modulation transfer functions in listeners with real and simulated hearing loss,"
J. Acoust. Soc. Am. 125(4), 2633-2633.

Book Chapters

Duchnowski, P., Lum, D.S., Krause, J.C., Sexton, M.G., Bratakos, M.S., and Braida., L.D. (2000). Development of Speechreading Supplements Based on Automatic Speech Recognition," Ch. 20 in Cued Speech and Cued Language for Deaf and Hard of Hearing Children, C. J. LaSasso, K. L. Crain, and J. Leybaert eds., Plural Publishing, San Diego CA., 467-486.

J. C. Krause, P. Duchnowski, L. D. Braida, "Automatic Cued Speech," Ch. 21 in Cued Speech and Cued Language for Deaf and Hard of Hearing Children, C. J. LaSasso, K. L. Crain, and J. Leybaert eds., Plural Publishing, San Diego CA., 487-502.

References:

Braida. L. D., Desloge, J. G., Reed, C. M., Delhorne, L. A. and Perez, Z. D. (2009).
"Forward-masking functions in listeners with real and simulated hearing loss,
J. Acoust. Soc. Am. 125, 2721

Desloge, J. G., Reed, C. M., Braida. L. D., Perez, Z. D., and Delhorne, L. A. (2009)
"Masking release for speech in listeners with real and simulated hearing impairment,"
J. Acoust. Soc. Am. 125, 2721.

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., Desloge, J. G., Perez, Z. D., Delhorne, L. A. and Braida. L. D. (2009)
"A comparison of gap-detection thresholds through audition and touch in listeners with real and simulated hearing impairment," J. Acoust. Soc. Am. 125,.

Rhebergen, K. S., Versfeld, N. J. and Dreschler, W. A. (2006).
"Extended speech intelligibility index for the prediction of the speech reception threshold in fluctuating noise ," J. Acoust. Soc. Am. 120(6) 3988-3997.