Hearing Aid Research

Sponsor
National Institutes of Health Grant R01 DC00117

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

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

Graduate Students
Alan Asbeck, Raymond Cheng, Joseph A. Frisbie, Raymond Goldsworthy, Ashish Koul, Paninah S. Rosengard

Technical And Support Staff
Ann K. Dix, Merry Brantley

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

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

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

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

Studies and Results

Characteristics of the Speech Signal

Our previous work has shown that sentences spoken “clearly" are more intelligible (roughly 17 percentage points) than those spoken “conversationally" for hearing-impaired listeners in a quiet background (Picheny et al., 1985) as well as for both normal hearing and hearing-impaired listeners in noise (Uchanski et al., 1996) and reverberation backgrounds (Payton et al., 1994). While talkers typically reduce their speaking rate when producing clear speech, we have also shown (Krause and Braida, 2002) that talkers can be trained to produce a form of clear speech at normal rates (clear/normal speech). This finding suggests that acoustical factors other than reduced speaking rate are responsible for the intelligibility advantage of clear speech. Our recent work has been aimed at determining which acoustical properties of clear/normal speech contribute most to its higher intelligibility through acoustic analysis of clear/normal speech and related signal processing transformations.

Initial acoustical measurements of clear/normal speech (Krause, 2001) have shown that only a subset of the acoustic properties of clear/slow speech were apparent in clear/normal speech. Moreover, some of the acoustic characteristics (e.g. segment duration, pitch, and voice-onset time) retained in clear/normal speech differed dramatically between talkers, suggesting that different talker strategies exist for producing clear speech at normal rates. In spite of these difficulties, three global acoustic properties associated with clear/normal speech were identified: increased energy near the second and third formants, higher average and greater range of F0, and increased modulation depth of low frequency modulations of the intensity envelope.

Signal processing transformations of conversational speech were developed in order to determine the relative contributions of each of these properties to the intelligibility advantage provided by clear/normal speech. The speech-based STI predicted that a majority of these processing schemes, presented singly and in combination, would improve intelligibility over conv/normal speech, when presented in wideband noise to normal hearing listeners. However, actual experiments with normal hearing listeners revealed an advantage only for clear/normal speech and formant processed speech. Moreover, hearing-impaired listeners did not obtain similar intelligibility benefits from clear/normal or formant processed speech as reliably as normal hearing listeners in noise, although these conditions did provide a statistically significant benefit for some individual hearing-impaired listeners and talkers (Krause, 2001).

One possible explanation of these results is that the benefits of clear/normal speech may be related to age, since the hearing-impaired participants in this study were older (40 to 65 years) than the normal hearing participants (19 to 43 years). Some studies report an age-related decline in speech reception for elderly listeners (Arlinger and Gustafsson, 1991), particularly those with hearing impairments (Hargus and Gordon-Salant, 1995). Another possibility is that the intelligibility benefits of these conditions do not extend to hearing-impaired listeners and that the additive noise model for simulating impairment in normal hearing listeners is inadequate. Although this simulation is appropriate for many mild to moderate impairments, it may not represent the effects of more severe impairments accurately. To investigate these two possibilities further, additional intelligibility tests will be conducted to evaluate the intelligibility of clear/normal, clear/slow, conv/normal, and conv/slow speech for young hearing-impaired, elderly hearing-impaired, and elderly normal-hearing listeners. These tests will differentiate the effect of age and impairment factors and clearly identify which groups can receive benefit from clear speech at normal speaking rates.

Because the intelligibility advantage provided by formant processing was not as large as the advantage provided by clear/normal speech, additional acoustic properties of clear/normal speech that contribute to its high intelligibility must exist. Analysis of three additional talkers who produced clear/normal speech after training (Krause and Braida, 2002) has begun that parallels measurements made previously both for the first two talkers (Krause, 2001) and for the acoustics of clear/slow speech (Picheny et al., 1986). In these studies, measurements were taken at three levels of detail: global, phonological, and phonetic. Global measurements (long-term spectra, temporal envelope modulations, pitch, and pause structure) have been completed, and no additional properties associated with clear/normal speech have been identified.

Another focus of our current work relates to the role of audibility of key words in the intelligibility advantage of clear speech. Although sentences were normalized for level in all previous work, individual words in a sentence are likely to have varied in level. The contribution of these level differences to intelligibility of key words is not fully understood. Uchanski et al. (1996) examined the intelligibility of key words in isolation, excised from clear and conversational speech, and found that the intelligibility advantage of clear speech was preserved. However, the individual excised words were not normalized beyond the normalization performed at the sentence level. We are currently repeating this experiment with normalization at the word level for clear/slow, clear/normal, and conv/normal speech. The relative intelligibility of key words normalized for level should provide a more precise indication of the maximum benefit that could possibly be achieved by signal transformations based on acoustic properties of clear speech.

In addition to this work, we plan to investigate the complexity of the speech database used for acoustic analysis. While a sentence database was appropriate for the intelligibility experiments, the primary problem with using a sentence database for the purposes of acoustic analysis is the presence of acoustic variability due to word positioning within sentences or phonetic context within words. For some acoustic properties, this variability could be large enough to mask the variability between conv/normal and clear/normal tokens. Therefore, after the current measurements are completed, a new database of conv/normal and clear/normal speech should be created that consists of sentences with a fixed number of phonetic contexts. This type of database would best satisfy the conflicting demands of acoustic analyses and intelligibility experiments. To capture various talker strategies in the new database, a large number of talkers should be recorded. An acoustic analysis of a database of this type is likely not only to identify additional acoustic properties associated with clear/normal speech but also to provide a comprehensive description of a variety of talker strategies. This information will be essential to the development of processing schemes that can provide robust intelligibility improvement for a variety of talkers and environments.

Computational Model of Speech Intelligibility

We continued our work to investigate a variety of methods to predict speech intelligibility for hearing-aid processed speech. To date we have focused on computing the Speech Transmission Index (STI) using speech as a probe stimulus. This year we focused on three techniques based on different ways to compute the Speech Modulation Transfer Function (SMTF): Houtgast and Steeneken's method (Houtgast et al. 1985) which we call the Houtgast Method, Drullman's phase-lock method (Drullman et al. 1994) referred to herein as the Drullman approach and an approach similar to Drullman's that is based on the Magnitude SMTF referred to herein as the Magnitude approach. In addition we have begun evaluating existing techniques used to assess speech quality of coded speech to see if there are tools in that field we can also apply to speech intelligibility prediction of hearing-aid-processed speech.

Payton, Chen and Braida (2002) compared the three different SMTF calculation techniques to theoretical MTF calculations for speech in noise and reverberation. When normal- hearing thresholds are assumed, the Drullman approach and the Magnitude SMTF are very similar and outperform the Houtgast method. Both methods almost always agree with theoretically calculated MTFs over the frequency range analyzed whereas the Houtgast method displays significant artifacts. In Payton, Chen and Braida (2002) we also considered how envelope spectra were affected by amplitude compression hearing aids. Two different hearing loss profiles were considered: A 50-dB flat loss and a moderate slopping loss where the processing conditions were the same as presented in Rosengard et al. (2001), i.e. linear amplification and multiple compression conditions. We treated the listener elevated thresholds as internal noise and then scaled the SMTFs accordingly. The linear amplification and the amplitude compression hearing aids were evaluated for both types of losses. All aids used the NAL-R frequency- gain prescription. For the flat-loss profile, we examined compression ratios of 2 and 3. For the slopping loss we used a compression ratio of 1.5 in the low frequency bands (<2.5 kHz) and 2.5 in the high frequency bands (>2.5 kHz). For flat loss profile, STMFs computed from all the three techniques were relatively flat and comparable at the low modulation frequencies. The differences appeared at high modulation frequencies. As has been noted previously, the STMF estimated from the Houtgast method increased with increasing modulation frequency, while STMFs computed from the other two methods dropped with increasing modulation frequency (Payton and Braida, 1999). It was also observed that at each modulation frequency, the SMTF from the Drullman approach was smaller than that computed by the Magnitude method. For slopping loss profiles, the phase-locked, Drullman, SMTF dropped to negative values at 10 Hz. It appeared that the output was exactly out of phase with the input at this modulation frequency. At frequencies above and below 10 Hz, the Drullman and Magnitude SMTFs were very similar (Payton et al. 2002).

We went on to carry the STI analysis on the hearing aid and hearing loss simulation data. Different SMTF techniques were used to compute Modulation Transfer Indices (MTIs) in each octave band. The IEC standard octave band weights and redundancy correction factors for a male talker were then applied to obtain the respective STIs (IEC 1998). Efforts had been focused on how well the STI techniques could predict listener intelligibility for hearing-aid processed speech. The relations between STI and intelligibility were examined. Similar results were observed for all the three techniques: STI values were higher in the linearly processed condition than those in the compression conditions although the subject scores were comparable in both cases. Within a class of processing (linear or amplitude compression), a reasonable consistency between STI prediction and the subjective intelligibility score is observed. The STI computed using the Drullman SMTF appears to have the tightest clustering for the amplitude- compression results but the data set is rather small to draw a firm conclusion (Payton, Braida et al. 2002).

The fact that STI approaches are not able to predict different processing classes (linear, amplitude compression) motivated us to explore new ways to improve the existing prediction algorithms. One choice is to imitate some aspects of human auditory perception into the prediction process, which is an area of great research interest to those in the speech coding community (e.g. Beerands and Stemerdink, 1992; Beerands Stemerdink, 1994; Voran, 1999; ITU-T Rec.P.862, 2001). One successful example is the PESQ (Perceptual Evaluation of Speech Quality) method, which applies the psychoacoustic models for evaluating narrowband telephone networks and speech codecs. PESQ has been accepted by ITU-T as an objective speech-quality measurement standard P.862. In brief, PESQ represents both the original and the degraded speech signals using psychophysical parameters. This is done by transforming the signals from the physical domain to the psychophysical domain through frequency warping and level compression. A distance measure is then applied to calculate the PESQ score. A conjecture is that we may use part of the PESQ idea to improve our current evaluation structure, or we may combine the STI approach and PESQ method to get better intelligibility prediction. We have already started the simulation applying the ITU-T P.862 standard to the hearing aid and hearing loss simulation data. Preliminary comparison between the PESQ score and the subject intelligibility score shows that the PESQ approach is good at distinguishing between different compression ratios for speech with no background noise but not for speech-in-noise conditions. Further investigation is ongoing.

Cochlear-Implant Research

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

The model of speech intelligibility under consideration is based on the speech transmission index (STI). STI was originally developed as a way of assessing room acoustics, and the original STI calculations are based on a system's response to specific test signals. Although these test signals are appropriate for assessing room acoustics, they are not suitable for many kinds of signal processing used in hearing aids and cochlear implants. As a result, several research groups, including our own, have attempted to develop methods for calculating STI based on the speech signal itself, rather than specific test signals.

We previously completed an analysis of the various methods for speech-based STI calculation described in the literature, establishing explicit relationships between the various speech-based STI calculations. This analysis revealed a number of issues that may hamper the performance of existing speech-based STI calculations for both noise reduction and cochlear-implant speech processing. We are currently developing and evaluating improved methods for speech-based STI calculation to address these issues.

In order to evaluate existing and novel methods for calculating speech-based STI, we are collecting intelligibility data under a variety of conditions. We plan a total of four experiments to address the following conditions: acoustic degradations (reverberation and additive noise), N of M processing (used in commercial cochlear-implant speech processors), spectral subtraction, and binaural noise reduction. Using normal-hearing subjects listening to a simulation of cochlear-implant speech processing, data collection has been completed for the first two experiments and is currently underway for the third experiment. We are currently recruiting cochlear-implant users to serve as subjects in all four experiments. Preliminary results demonstrate the shortcomings of existing methods for calculating speech-based STI; a novel method of speech-based STI shows promise where the conventional methods fail.

Signal Processing for Hearing Aids

Background noise presents a significant problem for hearing-aid users. Noise reduction algorithms attempt to preserve the desired sound source while reducing background noise. One common characteristic of many noise reduction algorithms is that they estimate parameters of the noise during the brief pauses that occur naturally in speech. In order to perform such estimation, the algorithm must first identify the intervals when the signal-to-noise (SNR) ratio is low. Our current work in this area is aimed at developing a simple method for estimating the range of the short-time SNR when desired speech is corrupted by background noise.

We propose a two-microphone system for the case where desired signal and noise arrive from different spatial locations. We assume that the desired signal arrives from a range of angles about array broadside and the noise arrives from a distinct range of angles. The microphone signals are first bandpass filtered, and then the intermicrophone correlation is computed and compared to a threshold to determine the range of the short-time SNR. System parameters include the center frequency and bandwidth of the bandpass filters and the threshold.

We have completed a theoretical analysis of this system in anechoic and reverberant environments. The analysis quantifies the effect of the parameters on system performance and indicates how to select the parameters in order to optimize performance in a range of acoustic environments. Computer simulations with broadband noise sources are in agreement with the theoretical results, and additional computer simulations with speech demonstrate the utility of this approach (Koul, 2003).

Characteristics of Sensorineural Hearing Impairment

A loss of cochlear compression may underlie many of the difficulties experienced by hearing- impaired listeners. An accurate and efficient behavioral estimate of basilar-membrane (BM) input- output functions may therefore be of clinical value. The magnitude of cochlear compression in human ears can be estimated using a forward-masking paradigm. This paradigm assumes that the response of the BM to tones around the characteristic frequency (CF) of the site of measurement is nonlinear, whereas the response to tones an octave or more below the CF is linear. Two procedures that use the forward-masking paradigm to derive BM input-output functions are growth-of-masking (GOM) (e.g., Oxenham et al., 1997) and temporal masking (TM) (e.g., Nelson et al., 2001). With the GOM procedure, masker levels necessary to just mask a signal are measured at several fixed levels. A high-pass noise is used to limit off-frequency listening. With the TM procedure, masker levels necessary to just mask a fixed low-level signal are measured as a function of the time delay between the masker and the signal. The use of a fixed low-level signal assures that the place along the BM of maximal signal excitation remains constant. Physiological studies in animals have shown BM response growth rates of around 0.2 to 0.3 dB/dB for sounds at or near CF (e.g., Ruggero et al., 1997).

This work was undertaken to determine the following: 1) Whether the two measures described above produce within-subject results that are consistent across a range of CFs and 2) which measure provides more efficient estimates of compression. GOM functions and TM curves were measured at signal frequencies (Fs) of 1, 2, and 4 kHz in a group of normal-hearing listeners for both an on- frequency (1.0 * Fs) and off- frequency (0.55 * Fs) tonal masker. The BM response (i.e., input- output function) at each CF was estimated by plotting masker levels for the off-frequency masker (output) against masker levels for the on- frequency masker (input), paired by signal level and signal delay for GOM and TM, respectively. The following function (Glasberg and Moore, 2000) was fit to the data to estimate compression:

(1) $$R = aL - bG_{\max} + V + c + d\left(G_{\max} - e\right)\left(1 - \frac{1}{1 + e^{f\left(L - L_0\right)}}\right)$$

where L is input level, Gmax is the maximum gain applied by the cochlear amplifier, and VS shifts the entire function along the ordinate. The value of Gmax and VS were allowed to vary as free parameters to achieve a least-squares fit to the data. The minimum of the first derivative of the fitted function was used to estimate the magnitude of compression. Interpretation of compression estimates from BM response functions depends on two assumptions: (1) that the BM response with masker level is linear for maskers well below signal frequency and (2) that the BM response to a masker at the signal place and time is constant for all maskers.

Compression values derived from GOM data and TM data were similar, both within and across subjects, at each signal frequency. However, compression values at lower signal frequencies (i.e., 1 and 2 kHz) were generally larger than expected for normal hearing listeners (i.e., > 0.3 dB/dB). This suggests that either the cochlea is less compressive at low CFs, or that one/both of the assumptions regarding interpretation of BM response functions are incorrect. In order to distinguish between these possibilities, BM response functions for the TM data were re-plotted using off-frequency masker levels for the 4 kHz signal condition as the output for all three signal conditions. This is justified if it is assumed that the decay of forward making is uniform across frequency. Slope estimates derived from these functions were all less than 0.3 dB/dB. This finding provides evidence for cochlear compression at both lower and higher CFs, and suggests compression at lower CFs is less frequency specific.

Test efficiency was evaluated by constructing 0.95 confidence intervals around the parameter Gmax, and hence minimum slope values, using the Jacobian matrix at the solution. A repeated- measures (fixed effects) analysis of variance showed that confidence intervals for GOM estimates were significantly smaller than those for TM estimates. Smaller confidence intervals for a given number of data points suggest that GOM is the more efficient measure of cochlear compression.

These data support the following conclusions: (1) The assumption that the BM responds linearly to an off-frequency masker may not be valid for low CFs. In the case of TM, BM response functions at low CFs can be estimated by comparing off-frequency masking levels at high frequencies to on-frequency masking levels at low frequencies. For the GOM procedure used here, it is not possible to circumvent the assumption of linearity for off-frequency maskers. (2) Compression values estimated from GOM data are larger than values estimated from TM data, particularly at low frequencies. However, at higher frequencies, where GOM and TM produce compression values that are more consistent within subjects as well as with physiological data in animals, GOM serves as a more efficient measure.

Currently, we are exploring these issues in listeners with cochlear- based hearing loss. Preliminary data from two hearing-impaired listeners with a moderately-severe flat loss suggest deriving BM response functions from TM curves may be problematic due to temporal processing deficits.

References

S. Arlinger, and H. A. Gustafsson, "Masking of speech by amplitude-modulated noise," *Journal of Sound and Vibration*, 15(3): 441-445 (1991).

J.G. Beerands, and J.A. Stemerdink, "A perceptual audio quality measure based on a psychoacoustic sound representation," *J. Aud. Eng. Soc.*, 40(12): 963-974 (1992).

J.G. Beerands, and J.A. Stemerdink, "A Perceptual Speech- Quality Measure Based on a Psychoacoustic Sound Representation," *J. Audio Eng. Soc.*, 42(3): 115-123 (1994).

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

B. R. Glasberg, and B.C.J. Moore, "Frequency selectivity as a function of level and frequency measured with uniformly exciting notched noise," *J. Acoust. Soc. Am.* 108, 2318-28 (2000).

S.E. Hargus, and S. Gordon-Salant, "Accuracy of Speech Intelligibility Index Predictions for Noise-Masked Young Listeners with Normal Hearing and for Elderly Listeners with Hearing" (1995).

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

IEC Sound System Equipment- Part 16: *Objective Rating of Speech Intelligibility by Speech Transmission Index*, 2nd Ed., International Standard, No. 60268-60216 (1998).

ITU-T Rec.P.862, "Perceptual evaluation of speech quality (PESQ): an objective method for end-to-end speech quality assessment of narrow- band telephone networks and speech codecs," *International Telecommunication Union*, Geneva, Switzerland (2001).

A. Koul, *Use of Intermicrophone Correlation in Estimating Signal to Noise Ratio*, M.E. thesis, MIT, 2003.

J.C. Krause, and L.D. Braida, "Investigating alternative forms of clear speech: The effects of speaking rate and speaking mode on intelligibility," *J. Acoust. Soc. Am.* 112(5): 2165-72 (2002).

J.C. Krause, *Properties of Naturally Produced Clear Speech at Normal Rates and Implications for Intelligibility Enhancement,* PhD dissertaion, MIT, 2001.

D.A., Nelson, A.C. Schroder, and M. Wojtczak, "A new procedure for measuring peripheral compression in normal-hearing and hearing- impaired listeners," *J. Acoust. Soc. Am.* 110: 2045-2064 (2001).

A.J. Oxenham, and C.J. Plack, "A behavioral measure of basilar membrane nonlinearity in listeners with normal and impaired hearing," *J. Acoust. Soc. Am.* 101: 3666-3675 (1997).

K.L. Payton, and L.D. Braida, "A method to determine the speech transmission index from speech waveforms," *J. Acoust. Soc. Am.*, 106: 3637-3648 (1999).

K. L. Payton, L.D. Braida, S. Chen, P. Rosengard, and R. Goldsworthy, "Computing the STI using speech as a probe stimulus," in *Past, Present and Future of the Speech Transmission Index*, Soesterberg, Netherlands, TNO Human Factors, 125-138 (2002).

K.L. Payton, S. Chen, and L.D. Braida, "Comparison of approaches to estimate the Speech Modulation Transfer Function," *J. Acoust. Soc. Am.* 111: 2431 (2002).

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

M.A. Picheny, N.I. Durlach, and L.D. Braida, "Speaking clearly for the hard of hearing I: Intelligibility differences between clear and conversational speech," *J. Speech Hear. Res.,* 28: 96-103 (1985).

P.F. Rosengard, L.D. Braida, and K.L. Payton, "Speech intelligibility and listener preferences for amplitude-compressed speech: Speech-based STI predictions and measurements on listeners with simulated hearing loss," *J. Acoust. Soc. Am.*, 109: 2355 (2001).

M.A. Ruggero, N.C. Rich, A. Recio, S.S. Narayan, and L. Robles, "Basilar-membrane responses to tones at the base of the chinchilla cochlea," *J. Acoust. Soc. Am.* 101: 2151-63 (1997).

R.M. Uchanski, S. Choi, L.D. Braida, C.M. Reed, and N.I. Durlach, "Speaking Clearly for the Hard of Hearing IV: Further Studies of the Role of Speaking Rate," *J. Speech and Hearing Res.* 39: 494-509 (1996).

S. Voran, "Objective estimation of perceived speech quality ---- part I: development of the measuring normalizing block technique," *IEEE Trans. Speech Audio Proc.* 7(4): 371-382 (1999).

PUBLICATIONS

Journal Articles

Published
J.C. Krause and L. D. Braida "Investigating alternative forms of clear speech: The effects of speaking rate and speaking mode on intelligibility," *J. Acoust. Soc. Am.* 112(5): 2165-2172 (2002).

Accepted for Publication
J.E. Greenberg, J.D. Desloge, and P.M. Zurek, "Evaluation of array-processing algorithms for a headband hearing aid," *Journal of the Acoustical Society of America*, March 2003.

Submitted for Publication
J. C. Krause and L. D. Braida, "Acoustic properties of naturally produced clear speech at normal speaking rates," submitted to *J. Acoust. Soc. Am.*, December, 2002.

Chapters in Books
K.L. Payton, L.D. Braida, S. Chen, P. Rosengard, and R. Goldsworthy, "Computing the STI using speech as a probe stimulus". in *Past, Present and Future of the Speech Transmission Index*, (Soesterberg, Netherlands, TNO Human Factors), 125-138 (2002).

Meeting Papers Presented
K.L. Payton, S. Chen, and L.D. Braida, "Comparison of approaches to estimate the Speech Modulation Transfer Function". *J. Acoust. Soc. Am.*, 111: 2431 (2002).

Theses
A. Koul, Use *of Intermicrophone Correlation in Estimating Signal to Noise Ratio*, M.E. Thesis, MIT, February 2003.