Hearing Aid Research

Professor Louis D. Braida, Dr. Paul Duchnowski, Dr. Julie Greenberg, Dr. Patrick M. Zurek, Dr. Kenneth W. Grant, Dr. Karen L. Payton, Dr. Christine M. Rankovic, Ann K. Dix, Merry Brantley, Phalguni P. Shah, Paninah S. Fine, Joseph A. Frisbie, Jean C. Krause, DinhYen Tran, Michael K. Qin

Specific Aims

Our long-term goal is to develop improved hearing aids for people suffering from sensorineural hearing impairments. 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:

1. To assess the relative contributions of various functional characteristics of hearing impairments to reduced speech-reception capacity.
2. To evaluate the effects of style of speech articulation and variability in speech production on speech reception by hearing impaired listeners.
3. To develop and evaluate analytical models that can predict the effects of a variety of alterations of the speech signal on intelligibility.
4. To develop and evaluate signal processing techniques that hold promise for increasing the effectiveness of hearing aids.

Studies and Results

Characteristics of Sensorineural Hearing Impairment

Listeners with severe hearing losses when stimulated at high sound levels are the main focus of this work. Ching et al. (1998) have shown that many hearing impaired listeners with severe or worse high-frequency losses achieve little or no benefit from audible speech in the high-frequency region. Hogan and Turner (1998) report similar results and suggest that if the loss exceeds 55 dB at 3 kHz and above there may be little benefit (or even a detriment) associated with making high-frequency speech cues audible.

We are studying such listeners using approaches based on both functional and analytical modelling. Ching et al. (1998) suggest that the relative impotence of high-frequency speech cues may result from perceptual distortions (reduced spectral and temporal resolution and increased spread of masking) that can be characterized by psychophysical measurements and incorporated in our functional models of hearing impairment. We are attempting to determine whether incorporating such distortions reduces the effectiveness of high frequency cues for listeners with normal hearing who experience the simulated losses.

The findings of Ching et al. (1998) and Hogan and Turner (1998) quantify deficits in effectiveness of high-frequency speech cues on a relative basis by comparison with data obtained from listeners with normal hearing tested under somewhat different presentation conditions. We are attempting to characterize these deficits in absolute terms using analytical models that we have used to predict the effects of combining speech cues from multiple sources (e.g., Braida, 1991; Braida, 1996).

Listeners will attempt to identify speech syllables presented in three conditions (e.g., 1) highpass filtered at 2800 Hz; bandpass filtered to 700-2800 Hz, and highpass filtered at 700 Hz -- the combination of the two previous bands). Ideal processor models will be used to predict the results observed in the combined-band condition from those observed in the sub-band conditions under the assumption of ideal cue integration (i.e. perfect integration with no interband interference).

Comparisons of scores and error patterns can distinguish the effects of poor in-band representation of high-frequency speech cues from those of perceptual interference between bands.

Characteristics of the Speech Signal

Clear Speech

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., 1995). While producing clear speech, however, talkers often significantly reduce their speaking rate. A more recent study (Krause and Braida, 1995) has shown 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 high intelligibility of clear speech. To gain insight into these factors, we are currently analyzing various signal processing transformations of conversational speech in order to determine which acoustical properties of clear/normal speech contribute most to its high intelligibility.

The signal transformations being pursued were derived from the results of a comprehensive acoustical analysis of the conversational (conv/normal) and clear/normal speech of two talkers from Krause's study (Krause, 1995). Acoustical measurements, as in an earlier study of the acoustics of clear speech by Picheny (1986), were taken at three levels of detail: global, phonological, and phonetic. Differences in acoustic characteristics of clear/slow speech relative to conv/normal speech were consistent with previously reported results. Many of these differences, however, were not apparent when comparing clear/normal speech to conv/normal speech. Since talkers presumably do not have time to retain all of the characteristics of clear/slow speech when speaking at normal rates, they must pick a subset of characteristics to continue to emphasize in clear/normal speech. The talkers in this study both showed increased modulation depth for slow modulation frequencies (less than 3Hz) and increased power near vowel formant frequencies, translating to a long-term spectrum with more relative power in frequencies above 1kHz. For other acoustic measurements (e.g. segment duration, pitch, and voice-onset time), however, the changes between clear/normal and conv/normal speech differed dramatically between talkers. These differences are most likely a result of different talker strategies for producing clear speech at normal rates. To investigate the possibility of different talker strategies, additional intelligibility tests were conducted in degraded listening environments other than additive noise. The results showed a significant talker effect for intelligibility scores, suggesting not only that talkers may have different strategies but also that some strategies may produce a type of clear/normal speech that is more robust in a variety of degraded environments.

Based on these results, the signal transformations currently under development are aimed at altering the following three properties of conv/normal speech: A) vowel formant energy will be increased by raising formant amplitudes and widening formant bandwidths, B) low-frequency modulations (<3--4Hz) will be enhanced, and C) F0 (pitch) average and range will be increased, since this acoustical property was exhibited by the talker whose clear/normal speech was most robust to other degradations. To assess the effectiveness of these intelligibility enhancement schemes, transformations will be evaluated singly and in combination. Both normal hearing and hearing-impaired listeners will participate in intelligibility experiments to evaluate whether listeners can derive intelligibility benefits from artificial manipulation of these acoustic properties.

Computational Model of Speech Intelligibility

In order to estimate the effects of a variety of types of signal processing transformations on speech, we are developing methods of predicting speech intelligibility from physical measurements made on speech waveforms. One method, the speech- based STI model (Payton and Braida, 1999), assumes that changes in envelope modulation spectra result in intelligibility changes.

Based on a review of the amplitude compression literature, we have implemented in software a general amplitude compression algorithm that allows us wide flexibility to vary the compression range and the attack and release times. We plan to conduct listening experiments using amplitude compressed speech to compare quantitatively experimental results with predictions based on our speech-based STI algorithm. Experimental compression parameters will be chosen based on published reports of relatively successful compression parameters.

We also plan to organize a formal comparison of our method of predicting speech intelligibility based on measurements of the speech waveform (Payton and Braida, 1999) with methods developed by others. This comparison will involve creating a database of speech test materials for which common intelligibility measures have been obtained and developing guidelines for the evaluation of intelligibility predictions.

Effects of Quasi-Periodic Interference

Recent investigations have found that digital wireless telephones can interfere wth hearing aid performance causing a degradation of communication for the hearing aid wearer.' This interference is the result of demodulation of the telephone's pulsed radio frequency (RF) transmission by non-linear elements in the hearing aid amplifier circuit. The resulting square wave and associated harmonics are amplified and spread throughout the hearing aid audio spectrum. The result is an undesired noise, that typically presents itself as a buzzing sound. Several groups have attempted to provide a standardized method for measuring the effects of this audio rectification type interference (ARTI) and establish levels of acceptability. ANSI C63.19 (ANSI, 1999) uses measures of perceived interference as the basis of its hearing aid ratings of immunity to RF emissions. Ratings based on predictions of the masking effects of ARTI noise on speech recognition may be another way to classify hearing aid performance. The speech Intelligibility Index (SII)3 is a computational method for predicting intelligibility based on measures of audibility. However, due to the quasi-periodic nature of ARTI noise it is not included in the scope of the Standard. The purpose of this study, conducted in cooperation with the City University of New York, was to quantify the effects of audio rectification type interference on speech recognition in normal hearing listeners and evaluate the ability of the STI (ANSI, 1997) to model these effects. Results suggest that there is a monotonic relationship between speech intelligibility and STI values for ARTI noise. This finding supports the use of the Sil in establishing standards of immunity for hearing aids to this class of interference.

Acoustic Correlates of Sound Quality Judgements

The long-term goal of this research is to develop a model for predicting subjective sound quality judgements based on acoustic measurements. Hearing-aid users often complain about the unpleasantness of sounds produced by their hearing aids in complex acoustic environments (for example, meetings, restaurants, and cocktail parties). Most previous hearing-aid research has focused on improving the objective intelligibility of speech, with significantly less attention paid to improving subjective sound quality. An objective method for predicting sound quality would improve the design and fitting of hearing aids by reducing the need for subjective evaluations, thereby making it more efficient to screen large numbers of candidate algorithms and evaluate large sets of parameter values.

Initial work in this area investigates the relationship between acoustic properties and sound quality ratings of speech in multi-talker babble. In a pilot study, stimuli are presented over headphones to normal-hearing listeners. The stimuli are presented at a variety of speech-to-noise ratios and speech levels. Listeners make paired-comparison judgements based on one of three sound quality aspects: background pleasantness, perceived speech intelligibility, and overall desirability.

Data from this preliminary study will be analyzed to investigate the effects of independent variables (subject, repetition, order of presentation) on the subjective sound quality judgements, as well as the correlation between acoustic variables (speech level, speech-to-babble ratio, overall sound level) and subjective sound quality judgements. This will help determine the feasibility of developing models for predicting subjective sound quality judgements based on acoustic measures. It will also contribute to efforts to predict one subjective sound quality judgement (overall level) based on constituent sound quality aspects (background pleasantness, perceived speech intelligibility).

Publications, References, and Theses

ANSI (1969). American National Standards Methods for the Calculation of the Articulation Index." S3.5-1969 (American National Standards Institute, New York).

ANSI (1997). American National Standards Methods for the Calculation of the Speech Intelligibility Index." S3.5-1997 (American National Standards Institute, New York).

ANSI (1999). American National Standards Methods of Measurement of Compatibility between Wireless Communications Devices and hearing Aids." C63.19 Draft Standard (American National Standards Institute, New York).

Braida, L. D. (1991). "Crossmodal integration in the identification of consonant segments," Q. J. Expt. Psych., 43A(3), 647-677.

Braida L. D., (1996). "Integration Models of Intelligibility," in @I{Speech Communication Metrics and Human Performance,} C.W. Nixon, Ed., USAF Armstrong Laboratory Technical Report 95-23.

Ching, T.Y.C., Dillon, H., and Byrne, D. (1998). "Speech recognition of hearing-impaired listeners: Predictions from audibility and the limited role of high-frequency amplification," J. Acoust. Soc. Am., 103, 1128-1140.

Hogan, C.A., and Turner, C.W. (1998). "High-frequency audibility: Benefits for hearing-impaired listeners," J. Acoust. Soc. Am. 104 (1) 432-441.

Krause, J.C. (1995). "The Effects of Speaking Rate and Speaking Mode on Intelligibility," S.M. Thesis, 1995.

Krause, J.C. and Braida, L.D. (1995). "The effects of speaking rate on the intelligibility of speech for various speaking modes," J. Acoust. Soc. Am. 98 (5), S2982.

Krause, J.C. and Braida, L.D. (1996). "Properties of naturally produced clear speech at normal speaking rates," J. Acoust. Soc. Am. 100 (4), S2828.

Payton, Karen L., Uchanski, Rosalie .M. and Braida, Louis D. (1995). "Intelligibility of conversational and clear speech in noise and reverberation for listeners with normal and impaired hearing," J. Acoust. Soc. Am. 95 (3), 1581-1592

Payton, Karen L. and Braida, Louis D. (1995). "Speech modulation transfer functions for different speaking styles," J. Acoust. Soc. Am. 98, 2982.

Payton, Karen L. and Braida, Louis D. (1999). "A method to determine the speech transmission index from speech waveforms," J. Acoust. Soc. Am. 106 (6) 3637-3648.

Payton, Karen L. and Braida, Louis D. (1995). "Speech modulation transfer functions for different speaking styles," J. Acoust. Soc. Am. 98 (5), 2982.

Picheny, M.A., Durlach, N.I., and Braida, L.D., (1985). "Speaking clearly for the hard of hearing I: Intelligibility Differences between Clear and Conversational Speech," J. Speech Hearing Res. 28, 96-103.

Picheny, M.A., Durlach, N.I., and Braida, L.D., (1986). "Speaking clearly for the hard of hearing II: Acoustic characteristics of clear and conversational speech", J. of Speech and Hearing Res. 29, 434-446.

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