Reprinted from Canadian Acoustics, Vol. 31 No. 3 (2003) pp 50-51

Developing a New Measure for Assessing Architectural Speech Security

J. S. Bradley and B. N. Gover
Institute of Research in Construction, National Research Council, Montreal Road, Ottawa, K1A 0R6, Canada

Introduction

This paper describes the process and problems of developing a new measure of architectural speech security. Such a measure is required to more accurately rate the probability of a listener outside a room being able to overhear conversations from within the room. Previous work has considered various levels of speech privacy, where some speech is intelligible [1]. However, speech security usually implies that none of the overheard speech is intelligible, or in some cases it is not even audible.

One can describe 3 levels of speech security. The first level would be when only a very small percentage or none of the overheard words are intelligible. Even when no words are intelligible, it is often still possible to recognize the rhythm or cadence of the speech. Finally, the highest level of speech security would be when all speech sounds from the adjacent space are completely inaudible.

Speech privacy and speech security have been related to signal/noise (S/N) type measures (where the signal is the speech from the adjacent space). The simplest would be a difference of A-weighted speech and noise levels. More sophisticated measures such as the Articulation Index (AI), and its more recent replacement the Speech Intelligibility Index (SII) (ANSI S3.5), are known to be better related to speech intelligibility within rooms. They more correctly weight the importance of S/N ratios at different frequencies and more accurately combine the various frequencies. These frequency weightings may not be optimum for speech security situations, and these measures are not ideal at very low levels of speech intelligibility.

Experimental Procedure

In this work, subjects rated simulated speech security situations. The subjects sat in a sound-isolated room and heard speech sounds, modified to simulate transmission through various walls, and combined with typical ambient noises. The speech and noise sounds were spatially separated and were precisely measured at the listener's head position in an acoustically dead environment.

Many variables will influence the intelligibility of overheard speech, including: talker gender and voice characteristics, speech material, voice level, wall transmission loss characteristics, ambient noise spectrum shape and level, listener hearing sensitivity, and other listener characteristics. Many of these were determined to be of less importance in pilot tests. Only subjects with negligible hearing loss were included.

In the main intelligibility experiment, 30 subjects each listened to 340 test sentences. The phonetically balanced and low predictability Harvard sentences were used [2], and 5 different sentences were used for each physical condition. Each condition was one of 68 combinations of varied: ambient noise, wall transmission loss characteristics, and S/N ratio. The conditions were chosen so that intelligibility ranged from 0 to 100%. A second experiment was intended to determine the thresholds of: (a) audibility of any speech sounds, (b) audibility of the cadence of the speech, and (c) the intelligibility of the speech. In this experiment the 20 best listeners from the first experiment each listened to 160 sentences. Again there were 5 sentences for each condition, and a range of ambient noise and wall transmission loss values. However, in this experiment conditions had, on average, much lower S/N values so that they included situations where no speech sounds were audible to the listeners.

Evaluation of Measures of Intelligibility

Figure 1(a) plots intelligibility scores versus measured SII values in the test sound fields. To simplify the plot, the results were averaged over all subjects. Although intelligibility scores increase with SII as expected, at SII=0 intelligibility is not zero. Therefore, SII (and AI) cannot be used to describe conditions for high levels of speech security which would correspond to acoustical conditions below SII=0, where SII is not defined. Figure 1(b) shows that differences of A-weighted levels are not limited in this way but are much less accurately related to intelligibility scores.

An example of a more successful measure is shown in Figure 1(c), which plots the same intelligibility scores versus a S/N ratio measure that included the same frequency weightings as the SII measure.

Speech Security Threshold Measurements

The 6 graphs of Figure 2 show the results of evaluations of the 3 types of thresholds and the effects of different weightings of the importance of each frequency band. Each graph shows the percentage of subjects with responses indicating: at least one word is intelligible (a) & (d), the cadence of the speech is audible (b) & (e), and some speech sounds are audible (c) & (f). In graphs (a)-(c) results are plotted against SII-weighted S/N ratios and in graphs (d)-(f) against LF-weighted S/N ratios, having greater emphasis on the lower frequencies.

If one considers the threshold to be when 10% of the subjects respond, the threshold for intelligibility is reached at -18.5 dB, for detection of cadence at -24 dB, and for audibility at -27 dB on the SII-weighted S/N ratio measure. Thus, complete speech security (inaudibility) requires speech levels to be about 8.5 dB lower than for the threshold of intelligibility. This would correspond to a

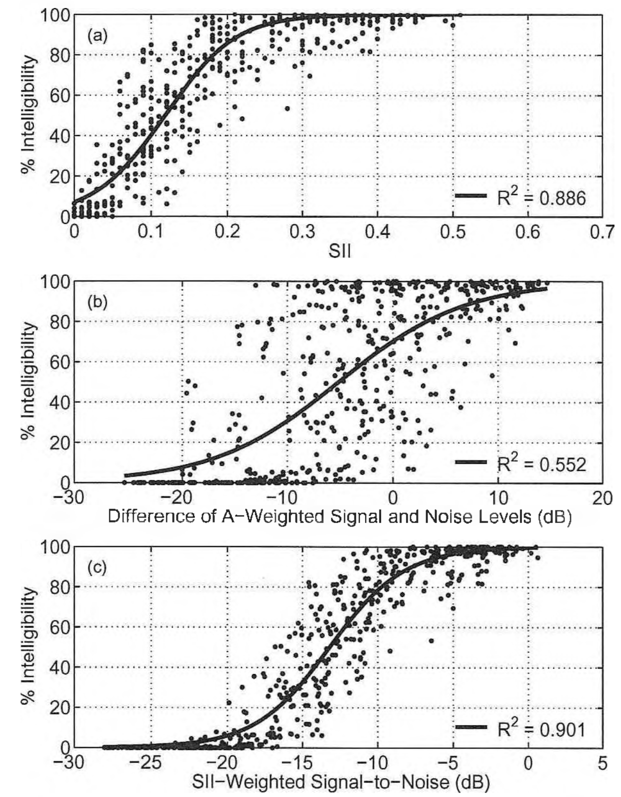

Fig 1. Speech intelligibility scores vs. (a) SII, (b) difference of A-weighted levels, and (c) SII-weighted S/N ratio.

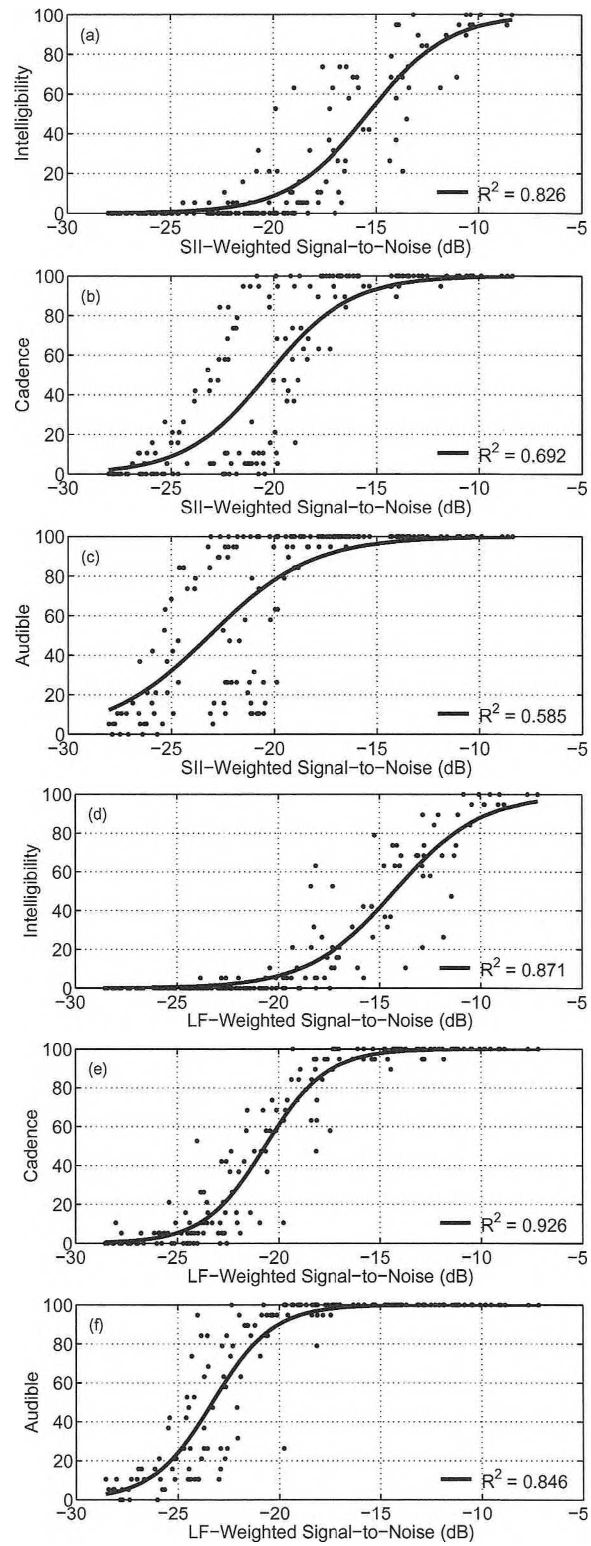

Fig 2. Percentage of subjects detecting some intelligibility (a) & (d), cadence (b) & (e), and audibility (c) & (f) vs. SII- weighted S/N (a)-(c), and LF-weighted S/N (d)-(f).

substantially better wall transmission loss. Although the SII-weighted S/N ratios predict the intelligibility threshold reasonably well, the LF-weighting provides more accurate estimates of the thresholds of cadence and audibility.

Conclusions

Existing measures of speech intelligibility and speech privacy are not adequate for evaluating the speech security of closed offices and meeting rooms.

The optimum frequency weighting for predicting the onset or threshold of intelligibility is different from that for predicting the threshold of the audibility or the cadence of speech sounds.

Speech security must be statistically described in terms of the percentage of listeners able to hear or understand speech from adjacent spaces.

Complete speech security, where speech sounds are totally inaudible, would require substantially better sound isolation of meeting rooms than is required for eliminating word intelligibility in adjacent spaces.

References

1. Cavanaugh et al., J. Acoust. Soc. Am. 34 (4), 475-492 (1962).

2. IEEE Trans. On Audio and Electroacoust., 227-246 (1969).