Proceedings of 20th International Congress on Acoustics, ICA 2010

23-27 August 2010, Sydney, Australia

A new system of speech privacy criteria in terms of Speech Privacy Class (SPC) values

J.S. Bradley and B.N Gover

Institute for Research in Construction, National Research Council, 1200 Montreal Rd., Ottawa, Canada, K1A 0R6
John.Bradley@nrc-cnrc.gc.ca

PACS: 43.55.HY

ABSTRACT

This paper describes a new system` of speech privacy criteria in terms of Speech Privacy Class (SPC) values. SPC values can be used to specify the required speech privacy for new construction or to assess the speech privacy of existing closed rooms. The ASTM E2638 measurement standard defines SPC as the sum of the measured average noise level at the position of a potential eavesdropper outside the room, and the measured average level difference between a source room average and the transmitted levels at the same potential eavesdropper location. For a given combination of level difference and ambient noise level, the likelihood of transmitted speech being audible or intelligible can be related to the probability of higher speech levels occurring in the meeting room, based on the statistics of speech levels from a large number of meetings. For a particular meeting room speech level, there is an SPC value for which transmitted speech would be at the threshold of intelligibility or even at the threshold of audibility. One can create a set of increasing SPC values corresponding to increasing speech privacy and for each SPC value one can give the probability of transmitted speech being either audible or intelligible. This makes it possible to accurately specify speech privacy criteria for meeting rooms and offices, varying from conditions of quite minimal to extremely high speech privacy, with an associated risk of a speech privacy lapse which is acceptable for each situation.

INTRODUCTION

This paper describes a new system of criteria for rating the speech privacy of closed rooms. Speech privacy is required so that eavesdroppers outside the room have difficulty understanding or in some cases even hearing speech from the room. The degree of speech privacy can vary from being able to understand some but not all of the words spoken in the room at points outside the room, to cases where it is very rarely possible to understand any of the words. It is also possible to have even higher privacy where it is difficult, or even impossible, to hear any speech sounds from the adjacent closed room. Very high speech privacy is often referred to as speech security.

Although it is often desirable to have some degree of speech privacy, very high privacy can be costly. Consequently, the amount of speech privacy must be designed to meet the needs of each particular situation. Usually the required degree of speech privacy is determined by the how sensitive the information is that is to be discussed in the room.

The likelihood of a speech privacy lapse can be described statistically and related to the probability of higher speech levels occurring in the closed room. In this paper, a system of speech privacy criteria is described that makes it possible to match the risk of a privacy lapse to the severity of the consequences of loss of information in each situation. Where more sensitive information is discussed, higher privacy is required to minimize the risk of loss of more critical information.

SPEECH PRIVACY BASICS

The intelligibility of speech increases with increasing speech-to-noise ratios at the position of the listener. Consequently, the speech privacy of closed rooms will increase with decreasing speech-to-noise ratios at the positions of potential eavesdroppers outside the room. There are many different ways to combine the influence of different frequencies in calculating signal-to-noise ratios, but our research [1] has shown that a uniform-weighted, frequency-averaged, signal-to-noise ratio over speech frequencies (SNR_{uni32}) best predicts the audibility and intelligibility of speech transmitted through various walls. SNR_{uni32} is given by,

$$SNR_{uni\,32} = \frac{1}{16}\sum_{f=160}^{5000}\{L_{ts}(f) - L_n(f)\}_{-32} \quad (1)$$

where, L_{ts} = transmitted speech level
L_n = ambient noise level
f is the $^1/_3$ octave band frequencies from 160 to 5000 Hz
-32 indicates that all $L_{ts}(f) - L_n(f)$ differences are clipped to never be less than -32, at which point speech would be inaudible.

Figure 1 illustrates a plot of speech intelligibility scores versus SNR_{uni32} values from the previous work [1].

The previous work also found SNR_{uni32} values corresponding to the thresholds of audibility and of intelligibility of transmitted speech sounds which are given in Table 1. These are the SNR_{uni32} values at which 50% of a panel of attentive listeners could just detect speech sounds or could just understand at least one word of short low predictability test sentences. These threshold values can be used to set design goals for particular situations.

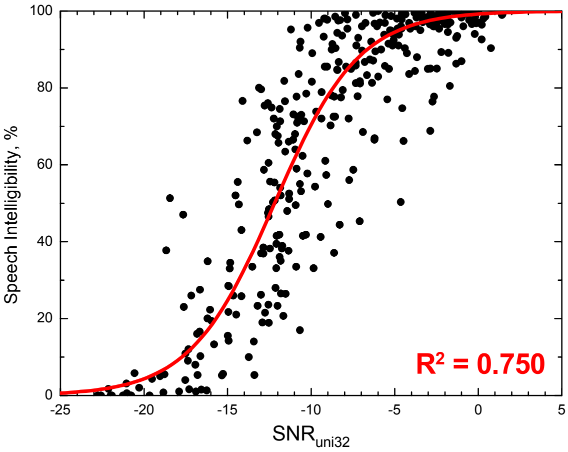

Figure 1. Speech intelligibility scores versus SNR_{uni32} values for speech sounds modified to simulate transmission through walls [1].

Table 1. Thresholds of Audibility and of Intelligibility of transmitted speech sounds [1].

SNR_{uni32}	Threshold
-16 dB	Intelligibility
-22 dB	Audibility

Subsequent work showed that, although the threshold of audibility was not affected, reflected sounds in rooms could affect the threshold of intelligibility [2]. However, these effects are not expected to be significant for most meeting room type spaces with reverberation times of 0.5 s or less. In more reverberant situations, the threshold of intelligibility can be increased a few dB.

In earlier speech privacy studies, the Articulation Index (AI) was used to rate the speech privacy of closed rooms [3]. Recently various speech privacy measures were compared [4], and the comparison of AI and SNR_{uni32} values is shown in Figure 2. These results suggest *Confidential Privacy* (AI = 0.05) is equivalent to an SNR_{uni32} value of about -14 dB. This would approximate the threshold of intelligibility in slightly reverberant environments [2]. Figure 2 also illustrates the limitation of AI values in that they approach asymptotically to 0 for low values indicative of high speech privacy. That is, AI values do not differentiate well among cases of high privacy and cannot be used to describe very high privacy where AI would be essentially zero.

ASTM E2638 MEASUREMENT STANDARD

To evaluate the speech privacy of a room we need to be able to measure SNR_{uni32} values at locations outside the room. A new procedure has been developed to do this and is described in the ASTM E2638 measurement standard [5]. The standard describes how to measure sound transmission from room average levels in the closed room to point receivers usually 0.25 m from the outside of the room in terms of frequency-averaged level differences (LD(avg)). Ambient noise levels are also measured at the same points outside the room in terms of frequency-averaged noise levels (L_n(avg)). In both cases "(avg)" indicates an arithmetic average over the speech frequency 1/3-octave band levels from 160 to 5000 Hz inclusive.

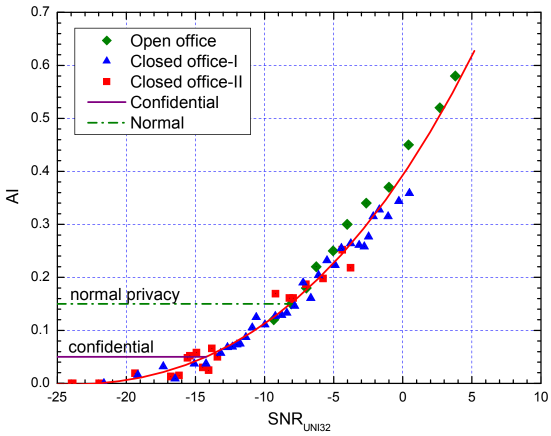

Figure 2. Plot of AI values versus SNR_{uni32} values for data from 3 previous studies. The horizontal solid and dash-dot lines indicate the *confidential (AI = 0.05)* and *normal (AI = 0.15)* speech privacy criteria respectively [4].

The speech privacy of a closed room will increase as both LD(avg) and L_n(avg) increase. The sum of these two quantities is referred to as the Speech Privacy Class (SPC) and can be used to rate the speech privacy of closed rooms.

$$SPC = LD(avg) + L_n(avg) \quad (2)$$

Conventional sound transmission measurements (e.g. ASTM E336, ISO140 Part V) assume diffuse sound fields in both spaces and measure the average transmission characteristics of the separating partition. Conventional transmission loss tests (illustrated in the upper part of Figure 3) are based on the measurement of room average levels in both adjacent spaces.

The new ASTM E2638 procedure measures level differences from a room average levels in the source room to spot receiver positions, usually 0.25 m from the outside of the meeting room (see lower part of Figure 3). A room average source level is used to represent the possibility of the talker being at any point in the room. This is achieved by measuring average test sound levels in the room using a combination of multiple source and microphone positions.

Spot receiver postions in the adjacent space represent a worst case scenario for speech privacy where an eavesdropper would be most effective if positioned close to the outside of the room. The ASTM E2638 procedure does not assume a diffuse field in the receiving space and produces measured level differences that will vary from point to point to indicate the likely variations in the speech privacy of the room. The receiver measurements at spot receiver positions close to the outer wall of the room are also little influenced by the acoustical properties of the adjacent space making it possible to measure into almost any adjacent space.

SPEECH LEVEL STATISTICS AND THE PROBABILITY OF A SPEECH PRIVACY LAPSE

For a given situation (i.e. for a particular combination of LD(avg) and L_n(avg) values), the likelihood of a speech privacy problem is related to the probability of higher speech levels occurring in the meeting room. If we can describe the statistical distribution of speech levels in typical meetings and meeting rooms, we can determine the probability of a speech privacy lapse in terms of the likelihood of speech levels exceeding either the threshold of audibility or the threshold of intelligibility at receiver positions in an adjacent space.

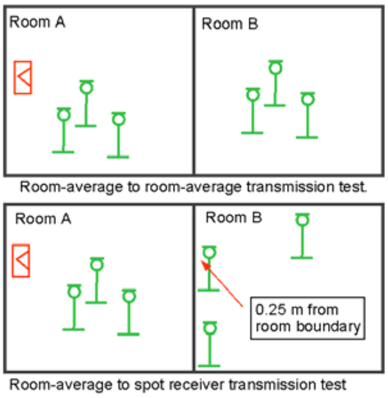

Figure 3. Comparison of ASTM E2638 method (lower) to that of conventional sound transmission measurements (upper). In both cases room average levels are measured in the source room (Room A). Although room average levels are also measured in the receiving space for conventional transmission tests (upper), the received levels are measured at spot receiver positions usually 0.25 m from the separating wall for the ASTM E2638 procedure (lower).

Information to describe the statistics of speech levels in meetings was obtained by placing data loggers around the periphery of meeting rooms for 24 hour periods. The data loggers recorded 10 s Leq values throughout each 24 hour period. The 10 s Leq values recorded during meetings were used to investigate speech levels in meeting rooms [6]. Table 2 gives a summary of the meetings and rooms measured. Few systematic effects of the variations in speech levels with properties of the rooms and their occupants were found.

Table 2. Summary of meeting rooms measured. (* includes 30 different rooms, 2 of which were measured with and without sound amplification systems).

Meeting and room parameters	Values
Number of meeting room cases* measured	32
Number of meetings measured	79
Number of people in each meeting	2 to 300 people
Range of room volumes	39 to 16,000 m^3
Range of room floor areas	15 to 570 m^2

In rooms with sound reinforcement systems, average levels were only about 2 dB higher than in rooms without sound amplification. The effect of sound reinforcement systems was minimal because speech levels were measured around the periphery of the rooms to represent speech levels incident on the room boundaries. This suggests that sound reinforcement systems were adjusted to provide levels, at more distant locations in larger rooms, that were similar to the speech levels found in smaller rooms without sound amplification.

Average meeting speech levels were found to increase systematically with ambient noise levels. The plot of increasing speech levels with increasing ambient noise levels in Figure 4 is an example of the Lombard effect [7]. Low ambient noise levels in meeting rooms are important for good intelligibility in the room, but also so that speech levels are lower and less likely to cause speech privacy problems at points outside the room.

The statistical characteristics of speech levels in meeting rooms were determined by creating a cumulative probability distribution plot of the 10 s Leq values of speech levels during all meetings. The distribution of all 110 773 Leq values is shown in Figure 5.

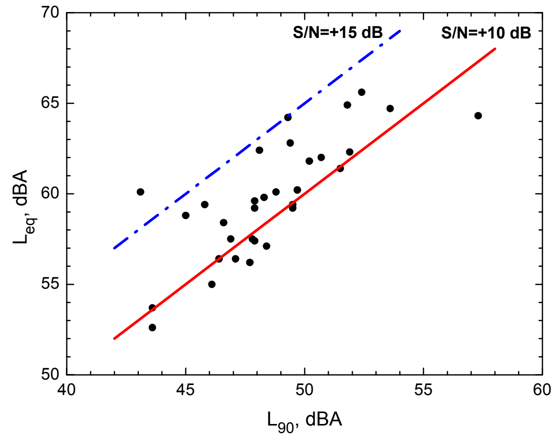

Figure 4. Meeting-average speech levels (L_{eq}) versus ambient noise levels in the meeting rooms (L_{90}). The solid diagonal line shows situations with a +10 dB speech-to-noise ratio and the dash-dotted line shows the more ideal conditions for good intelligibility of a +15 dB speech-to-noise ratio [6].

From the probabilities of the occurrence of various speech levels in Figure 5, one can calculate the corresponding average time interval between occurrences of particular speech levels taking into account the 10 s duration of each L_{eq} measurement of speech levels. Each probability indicates the frequency of occurrence of all speech levels up to and including the corresponding speech level on the x-axis. For example, a 90% probability corresponds to a speech level of 64.5 dBA, indicating that 90% of the time 10 s speech L_{eq} values would be no higher than 64.5 dBA. Hence, 10% of the time this speech level would be exceeded. There are 360 intervals of 10 s duration in one hour and this would correspond to speech levels exceeding 64.5 dB in 36 of them. On average there would be a 60 min / 36 = 1.67 minute interval between times when the 64.5 dBA speech level is exceeded.

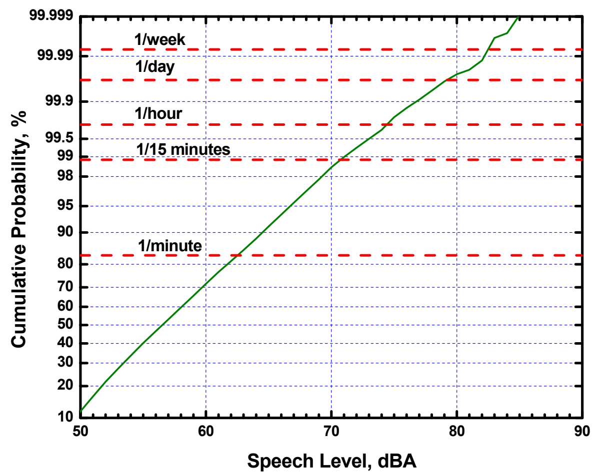

Figure 5. Cumulative probability distribution of 10 s speech L_{eq} values for the combined data from 79 meetings. The labels on the horizontal dashed lines (1/minute to 1/week) indicate the frequency of occurrence of the particular 10 s speech L_{eq} values.

SPEECH PRIVACY CLASS (SPC) CRITERIA

Speech privacy criteria can be given in terms of Speech Privacy Class (SPC) values (equation (2)). For each SPC value the probability of transmitted speech exceeding either the threshold of audibility or the threshold of intelligibility can be determined to describe the related risk of a privacy lapse. The audibility or intelligibility of speech can be related to the uniformly-weighted, frequency-averaged, signal-to-noise ratios (SNR_{uni32}), defined in equation (1). Table 1 gives SNR_{uni32} values for the thresholds of audibility and intelligibility of transmitted speech.

First we re-write equation (1) by replacing $L_{ts}(f)$ (the transmitted speech level) by $L_{sp}(f)$-LD(f) (the source room speech level less the measured level difference from the average level in the room to the level at a receiver outside the room).

$$SNR_{uni32} = \frac{1}{16}\sum_{f=160}^{5000}\{L_{sp}(f) - LD(avg) - L_n(f)\}_{-32} \qquad (3)$$

If we assume that the -32 dB clipping of the quantity in the curly brackets is usually not very important and can be neglected, then equation (3) can be simplified to equation (4).

$$SNR_{uni32} \approx L_{sp}(avg) - LD(avg) - L_n(avg) \qquad (4)$$

In equation (4) "(avg)" indicates arithmetic averaging of the 1/3-octave band values over the speech frequencies from 160 to 5000 Hz inclusive. This can be rearranged to the following,

$$LD(avg) + L_n(avg) \approx L_{sp} - SNR_{uni32} \qquad (5)$$

Finally, we usually want to design so that conditions meet or are below the threshold of intelligibility. From Table 1, this corresponds to an SNR_{uni32} of -16 dB or lower. The left side of equation (5), (LD(avg) + L_n(avg)) is the Speech Privacy Class (SPC). Substituting SNR_{uni32} = -16, we then have,

$$SPC \approx L_{sp} + 16 \qquad (6)$$

This tells us that for each SPC value there is a corresponding meeting room speech level that will produce conditions that just meet the threshold of intelligibility. Lower speech levels would not be expected to be intelligible at points outside the room. If the corresponding meeting room speech level in equation (6) is quite high, it will not occur very often and the room will have a reasonably high degree of speech privacy. Using Figure 5 we can say how often a particular speech level will occur and hence from equation (6) and knowledge of the SPC value, we can say how often speech transmitted from the room is likely to be intelligible. We could alternatively use the more stringent criterion for the threshold of audibility (SNR_{uni32} = -22 dB) and describe how often speech from the room would be just audible to an eavesdropper even though not intelligible.

Table 3. Summary of expected average time intervals between intelligibility and audibility lapses for Speech Privacy Class, *SPC*, values from 60 to 90.

SPC	Time between intelligibility lapses	Time between audibility lapses
60	0.32 min	-
65	0.76 min	-
70	2.87 min	0.62 min
75	18.03 min	2.09 min
80	2.28 hours	12.54 min
85	15.30 hours	1.53 hours
90	-	11.22 hours

Average expected intervals between intelligibility and audibility lapses were calculated for a range of SPC values [6] and are included in Table 3. Figure 5 includes horizontal dashed lines to indicate the speech levels corresponding to several time intervals (1/minute to 1/week).

SPC VALUES AND THEIR APPLICATION

Using the procedure described above, the risks of exceeding the thresholds of audibility and of intelligibility were determined for a range of SPC values. These are given for 5 point intervals of SPC values in Table 4. How often transmitted speech would be audible or intelligible is described in words that are explained in the legend below the table. It is seen that the 5 SPC values correspond to a wide range of conditions from quite minimal speech privacy to extremely high speech privacy.

In practice the middle 3 SPC values (75, 80 and 85) are probably of most practical use. Values of 90 and higher would correspond to essentially inaudible speech and lower values than 70 would suggest virtually no privacy at all. The 5 point SPC intervals represent a suitable perceptually small interval.

Speech privacy criteria would usually be determined by the most sensitive type of information to be discussed in the room. Proposed speech security criteria for use in Canadian federal government buildings would specify minimum SPC values of 75, 80 and 85 for rooms where Protected, Secret, and Top Secret information is to be discussed respectively. For more sensitive information, unique analyses would be required for each case.

Table 5 shows how the intermediate levels of privacy (SPC = 75, 80 and 85) relate to combinations of LD(avg) and L_n(avg). The three columns to the left of Table 5 give results for 3 different ambient noise levels referred to as "very quiet", "quiet" and "moderate noise". Ambient noise levels are given in terms of L_n(avg) values and are also converted to approximate A-weighted levels (L_n(A)). The conversion assumed a neutral noise spectrum decreasing at 5 dB per octave with increasing frequency. Below the ambient noise levels in Table 5, there are 3 rows of TL(avg) values (i.e. frequency-averaged transmission loss values). These have been empirically related to LD(avg) values [9],

$$TL(avg) \approx LD(avg)-1 \qquad (7)$$

This relationship makes it possible to estimate the sound isolation due to particular building elements from laboratory sound transmission loss test results. Finally, to the right of the TL(avg) values are the SPC values corresponding to the combination of the L_n(avg) values and the corresponding TL(avg) values in each row (as per equation (7)).

Table 4. Speech Privacy Classes (SPC) and the related risk of speech being audible or intelligible.

Category	SPC	Description
Minimal speech privacy	70	Frequently intelligible
Speech privacy	75	Occasionally intelligible, and frequently audible
Speech security	80	Very rarely intelligible, and occasionally audible
High speech security	85	Essentially not intelligible, and very rarely audible
Very high speech security	90	Unintelligible and essentially not audible

Legend
Frequently: about 1 per 2 minutes
Occasionally: about 1 per 15 minutes
Very rarely: about 4 per 8 hours
Essentially not: about 1 per 16 hours

The highlighted cells in Table 5 show the values of L_n(avg) = 24 dB and an as built TL(avg) = 55 combining to give an SPC = 80 which provides a high degree of speech privacy described as "Speech security". In table 4 this SPC value is described as corresponding to conditions where transmitted speech would be "Very rarely intelligible, and occasionally audible". From an analysis of the relationship between TL(avg) and STC values obtained from laboratory measurements of wood and light weight steel stud wall constructions, TL(avg) = 55 is approximately equal to an STC rating of 51. However, this is an approximate relationship with a standard deviation of ±4 points or more. It would also correspond to an R_w rating approximately the same as the STC value. These results suggest that with an as-built SPC rating of 80, quite high speech privacy can be achieved using relatively common constructions.

Of course the degree of speech privacy is also influenced by the ambient noise levels at the receiver position. In the above example a little higher noise level could provide very high speech privacy, but much quieter conditions would make it very difficult to achieve high speech privacy.

Table 5. Combinations of TL(avg) and L_n(avg) for some SPC values.

Ambient noise levels				
Very quiet	Quiet	Moderate noise		
14	24	34	⇦ L_n(avg)	
25	35	45	⇦ L_n(A)	
TL(avg) ≈ LD(avg)-1			SPC	Description
60	50	40	75	Speech privacy
65	55	45	80	Speech security
70	60	50	85	High speech security

For existing buildings it is usually possible to measure the actual ambient noise levels in spaces adjacent to meeting rooms. Such measurements should be over a long enough time interval to be able to indicate the lowest likely ambient levels when the room is in use. When lowest likely ambient noise levels cannot be measured, we can estimate them from previous measurements of noise levels in spaces adjacent to meeting rooms over 24 hour periods. When the lowest likely ambient noise level is taken to be the lowest 1 percentile level, the values shown in Table 6 were found for the day, evening and night periods [8].

Table 6. Estimates of lowest likely ambient noise levels in spaces adjacent to meeting rooms for 3 different times of day periods [8].

Period	Level, dBA
Day (8:00 to 17:00)	35
Evening (17:00 to 24:00)	30
Night (24:00 to 8:00)	25

CONCLUSIONS

The new SPC values provide a uniform system for rating all categories of speech privacy from very minimal privacy to extremely high speech security. SPC values can be measured to evaluate existing facilities or can be predicted for new facilities from laboratory tests of building elements. Of course to accurately predict the sound transmission from a meeting room to adjacent spaces in a real building, all sound paths must be considered. Flanking sound transmission via paths such as a common floor slab can severely limit the maximum possible sound isolation of a meeting room.

Although the procedures were developed for rating the speech privacy of meeting rooms, they could also be applied to other situations such as in health care facilities where speech privacy is often desired. To describe the risk of privacy problems in such other situations as health care facilities, it would be necessary to assess the probability of various speech levels occurring in those environments.

REFERENCES

1. Gover, B.N., and Bradley, J.S., "Measures for assessing architectural speech security (privacy) of closed offices and meeting rooms", *J. Acoust. Soc. Am*. vol. 116, (6) 3480-3490 (2004).
2. Bradley, J.S. and Apfel, M., and Gover, B.N., "Some Spatial and Temporal Effects on the Speech Privacy of Meeting Rooms", *J. Acoust. Soc. Am*., 125, (5) 3038-3051 (2009).
3. Cavanaugh, W.J., Farrell, W.R., Hirtle, P.W., and Watters, B.G., "Speech Privacy in Buildings", *J. Acoust. Soc. Am*. 34 (4), 475-492 (1962).
4. Bradley, J.S., "Comparisons of Speech Privacy Measures", *Proceedings Inter Noise 2009, Paper #191*, (Ottawa, August 2009).
5. ASTM E2638-08, "Standard Method for Objective Measurement of the Speech Privacy Provided by a Closed Room", *ASTM International* (formerly American Society for Testing and Materials), West Conshohocken, PA, U.S.A.
6. Bradley, J.S. and Gover, B.N., "Speech Levels in Meeting Rooms and the Probability of Speech Privacy Problems", *J. Acoust. Soc. Am*. 127 (2) 815-822(2010).
7. Junqua, J.C., "The influence of acoustics on speech production: A noise-induced stress phenomenon known as the Lombard reflex," *Speech Commun*. 20, 13-22 (1996).
8. Bradley, J.S. and Gover, B.N., "Speech and Noise Levels Associated with Meeting Rooms", *NRC-IRC Research Report RR-170* (2004). http://www.nrc-cnrc.gc.ca/obj/irc/doc/pubs/rr/rr170/rr170.pdf
9. Bradley, J.S. and Gover, B.N., "Validation of Architectural Speech Security Results", IRC Research Report, RR-221, March 2006. URL: http://www.nrc-cnrc.gc.ca/obj/irc/doc/pubs/rr/rr221/rr221.pdf