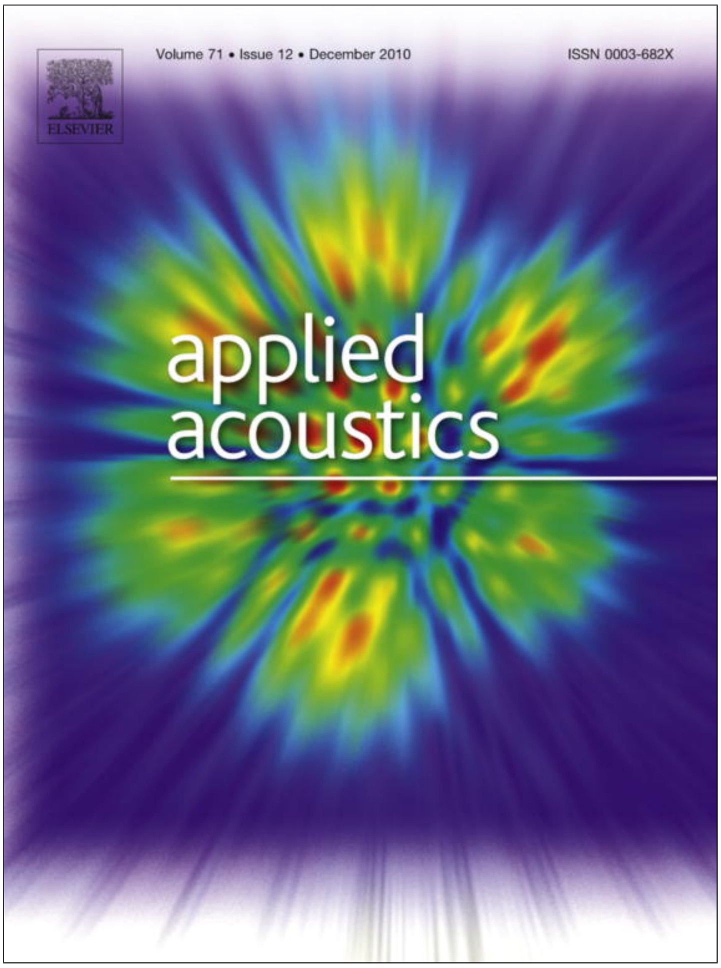
Volume 71 • Issue 12 • December 2010
ISSN 0003-682X
ELSEVIER
applied acoustics

Applied Acoustics 71 (2010) 1156–1161

Contents lists available at ScienceDirect

Applied Acoustics

journal homepage: www.elsevier.com/locate/apacoust

Verbal communication and noise in eating establishments

Jens Holger Rindel *

Multiconsult AS, Post Box 265, Skøyen, N-0213 Oslo, Norway

ARTICLE INFO

Article history:
Received 25 March 2010
Received in revised form 24 June 2010
Accepted 15 July 2010
Available online 11 August 2010

Keywords:
Room acoustics
Noise
Speech level
Lombard effect
Restaurants

ABSTRACT

A new simple prediction model has been derived for the average A-weighted noise level due to many people speaking in a room with assumed diffuse sound field. Due to the feed-back influence of noise on the speech level (the Lombard effect), the speech level increases in noisy environments, and the suggested prediction model gives a 6 dB reduction of the noise level by doubling the equivalent absorption area of the room. This is in contrast to the lowering by 3 dB by doubling of the absorption area for a constant power sound source. The prediction model is verified by experimental data found in the literature. In order to achieve acceptable conditions for speech communication within a small group of people, a guide for the recommended minimum absorption area per person in eating establishments is provided.

1. Introduction

Noise from people speaking in restaurants, canteens and at social gatherings in a reverberant environment is often a nuisance because it can be very loud, and a conversation may only be possible with a raised voice level. Because of the noise and the difficulties associated with a conversation the visitors may leave the place with a feeling of exhaustion.

It is a well known phenomenon that many people speaking in a room can create an increasing sound level, because the ambient noise from the other persons speaking means that everyone raises the voice, which again leads to a higher ambient noise level. This effect is called the Lombard effect after the French otolaryngologist Étienne Lombard (1869–1920), who as early as 1909 was the first to observe and report that persons with normal hearing raised their voice when subjected to noise. The average relationship between speech level and ambient noise level (the so-called Lombard slope) has been studied in an overview paper by Lazarus [1] and laid down in ISO 9921 [2].

Tang et al. [3] suggested a prediction model for noise in an occupied enclosure with repeated iterations by assuming a raised voice level due to the ambient noise, which again increases due to the raised voice level. Measurements in a canteen were also reported, with number of occupants varying from very few and up to ca. 300 while the measured A-weighted sound pressure levels varied from 57 dB to 75 dB. They applied the absorption of 0.44 m^2 per person, but the assumed absorption per person was found to have very little influence on the predicted noise level.

Kang [4] used a computer model and the radiosity method to predict sound pressure levels in dining spaces. A constant sound power from all speakers was assumed. A parametric study was carried out to examine the basic characteristics of conversation intelligibility in dining spaces and to study the effect of increasing absorption area per person, ceiling height, etc.

Navarro and Pimentel [5] reported the relationship between number of people and the measured sound pressure level due to the noise from speech in two large food courts. In one foot court the measured A-weighted sound pressure level was up to 74 dB with ca. 345 people, whereas up to 80 dB was measured in the other foot court with ca. 540 people. Attempts to explain the results by a simplified analytical model showed some similarities with the measured results assuming raised vocal effort and an average group size of either 2 or 4 people per talker.

Hodgson et al. [6] measured noise levels in 10 eating establishments and reported A-weighted sound pressure levels between 45 and 82 dB. They also described an iterative model for predicting the noise levels including the Lombard effect. Using an optimisation technique they found the best estimates for some unknown parameters in the model, e.g. that absorption per person varied between 0.1 and 1 m^2 with an average of 0.5 m^2, the Lombard slope was on average 0.69 dB/dB, and the average number of customers per talker was about 3.

The aims of this paper are to present a simple prediction model that takes the Lombard effect into account and to try to answer the following two questions:

* Tel.: +47 21 58 53 29; fax: +47 21 58 50 01.
E-mail address: jens.holger.rindel@multiconsult.no

0003-682X/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.apacoust.2010.07.005

- How much will the sound pressure level of the noise from speaking persons in an eating establishment be reduced by acoustic treatment if the absorption area is increased by a certain amount, say by a factor of 2?
- Is it possible to give design guidelines for satisfactory or good acoustic conditions in eating establishments?

2. Vocal effort in background noise

The vocal effort is characterised by the equivalent continuous A-weighted sound pressure level of the direct sound in front of a male speaker in a distance of 1 m from the mouth. The symbol is $L_{S,A,1m}$ and a description of the vocal effort in steps of 6 dB are given in Table 1. Thus normal vocal effort corresponds to a sound pressure level between 60 and 65 dB in the distance of 1 m. Speech at levels above 75 dB may be more difficult to understand than speech at lower vocal effort [2].

The Lombard effect means that $L_{S,A,1m}$ increases as a function of the A-weighted ambient noise level $L_{N,A}$. This can be described by the rate c (the Lombard slope). Lazarus [1] made a review of a large number of investigations, and he found that the Lombard slope could vary in the range $c = 0.5–0.7$ dB/dB. The Lombard effect was found to start at an ambient noise level around $L_{N,A} = 45$ dB and a speech level $L_{S,A,1m} = 55$ dB [1]. Assuming a simple linear relationship for ambient noise levels above 45 dB, the speech level can be expressed in the equation:

$$L_{S,A,1m} = 55 + c \cdot (L_{N,A} - 45)(\mathrm{dB}) \tag{1}$$

for $L_{N,A} > 45$ dB. In a free sound field the relation between sound power level $L_{W,A}$ and sound pressure level in the distance r is (see e.g. [7] Chapter 3):

$$L_{S,A,r} = L_{W,A} + 10\log Q - 10\log(4\pi r^2)(\mathrm{dB}) \tag{2}$$

where Q is the directivity factor and r is the distance (in m) from the sound source. For the talker as a sound source it is assumed as a rough approximation that $Q = 2$, and referring to the sound pressure level in a distance $r = 1$ m, the A-weighted sound power level for one person speaking is:

$$L_{W,A} = L_{S,A,1m} + 8\ \mathrm{dB} \tag{3}$$

The same relationship was used by Hodgson et al. [6].

3. Suggested prediction model

In order to estimate the sound pressure level from noise sources in a room, the simplest assumption is that of a diffuse sound field. If there are N_S noise sources active at the same time and the sound power level of each is $L_{W,A}$, the average sound pressure level in the room, i.e. the ambient noise from N_S speaking persons is calculated by (see e.g. [7] Chapter 3):

$$L_{N,A} = L_{W,A} + 10\log N_S - 10\log\left(\frac{A}{4}\right)(\mathrm{dB}) \tag{4}$$

where A is the equivalent absorption area (in m^2) of the room. This equation may be sufficiently accurate in rooms with a high ceiling and little absorption. However, in rooms with a low ceiling and high absorption the sound pressure level decreases with the distance from the source, and the average sound pressure level may be lower than predicted with Eq. (4).

Insertion of (1) and (3) in (4) yields:

$$L_{N,A} = \frac{1}{1-c}\cdot\left(69 - c\cdot 45 - 10\log\left(\frac{A}{N_S}\right)\right)(\mathrm{dB}) \tag{5}$$

This equation gives the ambient noise level that can be estimated from the equivalent absorption area of the room and the number of people speaking. If the room has the volume V (m^3), if the reverberation time in unoccupied state is T_0 (s), and if a diffuse sound field can be assumed, the Sabine equation gives the following estimate of the equivalent absorption area including the contribution from N persons (see e.g. [7] Chapter 3):

$$A = \frac{0.16\cdot V}{T_0} + A_p \cdot N(\mathrm{m}^2) \tag{6}$$

where A_p is the sound absorption per person in m^2. Introducing the group size, here defined as the average number of people per speaking person, $g = N/N_S$, insertion of (6) in (5) yields:

$$L_{N,A} = \frac{1}{1-c}\cdot\left(69 - c\cdot 45 - 10\log\left(g\cdot\left(\frac{0.16\cdot V}{T_0\cdot N} + A_p\right)\right)\right)(\mathrm{dB}) \tag{7}$$

This equation is the suggested prediction model for the noise that can be estimated in an eating establishment with the main input parameters: room volume, reverberation time in empty state, and number of people in the room. Appropriate values for the three parameters c, g, and A_p will be discussed in the following. From Eq. (7) it appears that the Lombard slope c and the group size g may have a strong influence, whereas the sound absorption per person A_p may be of minor importance if small compared to the equivalent absorption area per person in the room. In restaurants where people are seated at tables it is assumed that the absorption per person is less than for standing persons. Actually, in case of seated persons A_p is the additional amount of absorption compared to the empty chairs. Also the amount of clothes may vary according to the actual climatic conditions. For seated persons in light summer clothes it is suggested that A_p may be as low as 0.2 m^2.

4. Verification of the prediction model

Two examples of measured noise levels as a function of number of people in food courts are reported by Navarro and Pimentel [5]. The Food Court J had a volume of 7228 m^3, and reverberation time 1.3 s at mid frequencies (500–1000 Hz), while Food Court L had a volume of 3133 m^3, and reverberation time 0.9 s at mid frequencies. In both cases measurements of the A-weighted sound pressure level were reported with different number of people present. However, it is not known how many people were actually speaking. In Figs. 1 and 2 the reported results from the two food courts are compared with the prediction model (7) using the Lombard slope $c = 0.5$ dB/dB, the sound absorption per person $A_p = 0.2$ m^2, and different values of the group size g (between 2 and 4). In both cases it is found that a group size of 3 gives the best overall agreement of the model with the measured results.

Another set of measured data is reported by Tang et al. [3]. The noise level was measured continuously in a canteen for 2.5 h during lunch time, where the number of people increased in the first hour from nil to around 250 (Measurement A in Fig. 3). In the later 1.5 h the number of people gradually decreased, but the noise level did not decrease as much as could be expected, and at the end of the measurements around 50 people were left, but the noise level was about 5 dB higher than with the same number of people at the

Table 1
Description of vocal effort at various speech levels, after ISO 9921 [2].

$L_{S,A,1m}$ (dB)	Vocal effort
54	Relaxed
60	Normal
66	Raised
72	Loud
78	Very loud

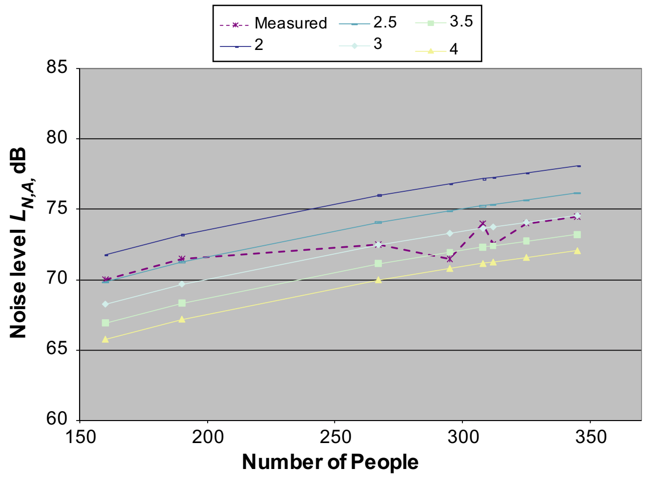

Fig. 1. Measured and predicted noise level for Food Court J [5] (7228 m^3) as a function of the number of people present. The parameter on the predicted curves is the assumed group size, *g*.

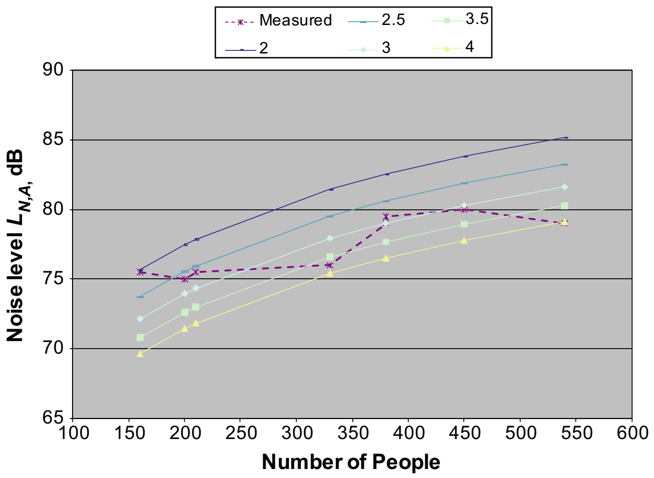

Fig. 2. Measured and predicted noise level for Food Court L [5] (3133 m^3) as a function of the number of people present. The parameter on the predicted curves is the assumed group size, *g*.

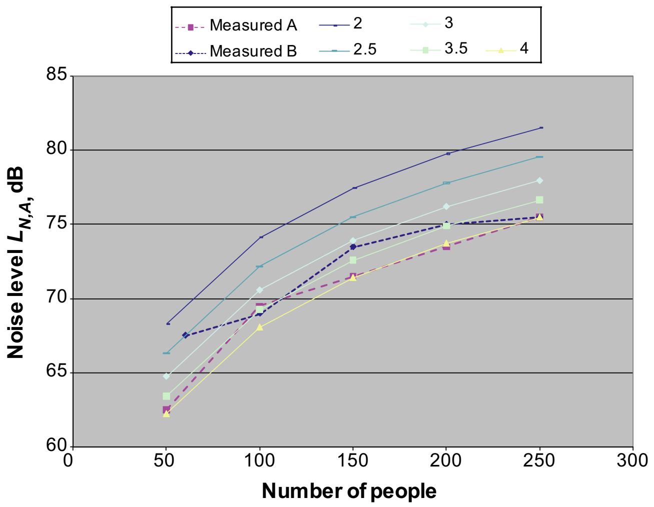

Fig. 3. Measured and predicted noise level for a canteen [3] (1235 m^3) as a function of the number of people present. Measurement A: first period with increasing number of people; Measurement B: second period with decreasing number of people. The parameter on the predicted curves is the assumed group size, *g*.

beginning. The canteen had a volume of 1235 m^3, and the unoccupied reverberation time 0.47 s at mid frequencies. The measured results are compared with the prediction model (7) using the Lombard slope $c = 0.5$ dB/dB, the sound absorption per person $A_p = 0.2$ m^2, and different values of the group size *g*. The best overall agreement with the prediction model is obtained with a group size of 3.5. However, in Measurement A between 150 and 250 people, a very good agreement is obtained with a group size of 4, indicating that people are not talking so much in the beginning of the lunch, whereas the later part of the lunch represented by Measurement B matches better with a group size of 3, i.e. more people talking. Thus it is clear that the group size *g* should not be considered constant, but varies according to the social character of the gathering.

Other values of the Lombard slope has been tried, but only for $c = 0.5$ dB/dB the average deviation from the measured data is acceptable small (see Table 2). A small variation of ±0.1 dB/dB leads to deviations of 4–8 dB from the measured data. So, it is concluded that for the cases analyzed here, the Lombard slope should be fixed at $c = 0.5$ dB/dB.

Table 3 shows the influence of small variations in group size and sound absorption per person. It is seen that a variation in A_p

Table 2
Deviation between predicted and measured noise levels in three eating establishments, using $A_p = 0.2$ m^2 and different values of the parameters *c* and *g*.

c, dB/dB	Average deviation, dB Group size, *g*				
	2	2.5	3	3.5	4
Food Court J; Navarro and Pimentel [5]					
0.4	−1.7	−3.3	−4.7	−5.8	−6.7
0.5	3.4	1.5	**−0.1**	−1.4	−2.6
0.6	11.2	8.8	6.8	5.1	3.7
0.7	24.1	20.9	18.2	16.0	14.1
Food Court L; Navarro and Pimentel [5]					
0.4	−2.6	−4.2	−5.5	−6.6	−7.6
0.5	3.4	1.4	**−0.1**	−1.5	−2.6
0.6	12.3	9.9	7.9	6.3	4.8
0.7	27.3	24.0	21.4	19.2	17.2
Canteen; Tang et al. [3]					
0.4	−0.2	−1.8	−3.1	−4.2	−5.2
0.5	5.0	3.0	1.4	**0.1**	−1.1
0.6	12.7	10.2	8.3	6.6	5.1
0.7	25.5	22.3	19.6	17.4	15.5

Table 3
Deviation between predicted and measured noise levels in three eating establishments, using $c = 0.5$ dB/dB and different values of the parameters A_p and *g*.

A_p (m^2)	Average deviation, dB				
	Group size, *g*				
	2	2.5	3	3.5	4
Food Court J; Navarro and Pimentel [5]					
0.2	3.4	1.5	**−0.1**	−1.4	−2.6
0.3	3.2	1.3	−0.3	−1.7	−2.8
0.4	3.0	1.0	−0.6	−1.9	−3.1
0.5	2.7	0.8	−0.8	−2.1	−3.3
Food Court L; Navarro and Pimentel [5]					
0.2	3.4	1.4	**−0.1**	−1.5	−2.6
0.3	2.9	1.0	−0.6	−1.9	−3.1
0.4	2.5	0.6	−1.0	−2.3	−3.5
0.5	2.1	0.2	−1.4	−2.7	−3.9
Canteen; Tang et al. [3]					
0.2	5.0	3.0	1.4	**0.1**	−1.1
0.3	4.7	2.8	1.2	−0.2	−1.3
0.4	4.4	2.5	0.9	−0.4	−1.6
0.5	4.2	2.3	0.7	−0.7	−1.8

from 0.2 to 0.5 m^2 has a rather small influence on the results (less than 1 dB), and it has the same effect as a reduction of g by 0.25. So, in case 1 the match is equally good with the combination A_p = 0.5 m^2 and g = 2.75 as with A_p = 0.2 m^2 and g = 3. In case 3 the match is equally good with the combination A_p = 0.5 m^2 and g = 3.25 as with A_p = 0.2 m^2 and g = 3.5.

It appears from the above examples that the suggested prediction model may give the best results within a range of ±2 dB using the following values of the parameters:

- c = 0.5 dB/dB (Lombard slope).
- A_p = 0.2–0.5 m^2 (additional absorption per person compared to that of empty chairs, also depending on amount of clothing).
- $g = N/N_S$ = 3 (may vary between 2 and 4 in typical eating courts and canteens, but higher values may occur in places with little verbal communication).

The average group size of 3 for eating establishments was also that suggested by Hodgson et al. [6]. The verification examples above are with people seated at tables; the model has not been tested against examples with standing people where a somewhat higher absorption per person may be expected. However, the prediction model has little sensitivity to the absorption per person in the examples above.

5. Discussion

The suggested prediction model (7) has been verified to be in reasonably good agreement with measured data reported by other researchers from eating establishments with a variation of number of people in a range from ca. 50 to 540 people. However, it is obvious that in general noise from speech in such cases cannot be predicted with a very high accuracy, simply because there are unknown parameters related to how much people actually want to talk. This may depend on the type of gathering, which can be more or less lively, how well people know each other, the age of the people, and other social circumstances.

5.1. The group size

In the suggested prediction model the most difficult parameter to estimate for a specific case is the average group size g, which can also be looked at as a measure of the arousal level of the people gathered. A higher number (e.g. 4) indicates a relatively quiet party, whereas a low number (e.g. 2) suggests a very lively party. If the party is like a reception and the people are served alcohol, the arousal level may increase with time, meaning a decreasing group size and thus a further increasing noise level with time.

5.2. Ten test cases

As a test of the suggested prediction model the 10 eating establishments reported by Hodgson et al. [6] are considered. The places belong to four different categories: cafeterias (C1–C2), bistros (B1–B3), restaurants (R1–R3), and senior residence dining rooms (S1–S2). As input data are used the volume, the unoccupied reverberation time (RT) at mid frequencies (500–1000 Hz), and the number of seats, which is assumed to equal the number of people when fully occupied, see Table 4. The sound absorption per person is here estimated to be A_p = 0.5 m^2 in all 10 cases. For each eating establishment the A-weighted sound pressure level has been monitored over one day of normal operation, and the highest level in the reported range has been used for comparison with the present prediction model, using a fixed Lombard slope of 0.5 dB/dB and adjusting the group size to give results that fit within ±0.5 dB, see Table 4. This optimised group size is seen to vary between a minimum of 2.5 in case B3 and a maximum of 9 in case S1.

For the bistros, cafeterias and restaurants the average group size found by this method is around 4, i.e. a little higher than the group size of 3 found earlier. This may be explained by the fact that we are now using the number of seats instead of the number of persons; so the actual number of persons is not known and it may be somewhat less than the number of seats. In other words, if the number of persons is slightly overestimated, this is compensated for by assuming a bigger group size. The two senior residence dining rooms are very different from the other cases, and they may be characterized by a group size as high as 8, i.e. the conversation here is not as lively as in the other eating establishments.

The last two columns in Table 4 show the results using a group size of 4 in the eight cases of bistros, cafeterias and restaurants, and a group size of 8 in the two cases of senior residence dining rooms. The deviations from the measured noise levels are within ±1 dB in eight of the 10 cases; only C2 and B3 show larger deviations within ±4 dB.

5.3. The acoustic treatment

Concerning the efficiency of acoustic treatment of a room, the ambient noise level due to N_S speaking persons is estimated from (5) with the insertion of the suggested Lombard slope c = 0.5 dB/dB:

$$L_{N,A} = 93 - 20\log\left(\frac{A}{N_S}\right)(\text{dB}) \tag{8}$$

Thus, an increase of the absorption area per speaking person by a factor of 2 leads to a reduction of the ambient noise level due to speech by 6 dB. This finding is in contrast to the lowering by 3 dB per doubling of the absorption area for a sound source with

Table 4
The results obtained with the suggested prediction model on 10 eating establishments measured and reported by Hodgson et al. [6].

EE	Volume (m^3)	RT unocc. (s)	No. of seats	Measured $L_{N,A}$ (dB)	Lombard slope, c	A_p (m^2)	Abs. area (m^2)	Optimised group size, g	Assumed group size, g	Calculated $L_{N,A}$ (dB)	Deviation (dB)
C1	619	0.5	120	75	0.5	0.5	258	3.5	4	**74.3**	−0.7
C2	412	1.0	100	76	0.5	0.5	116	6	4	**79.7**	3.7
B1	692	1.5	72	77	0.5	0.5	110	4	4	**77.3**	0.3
B2	384	1.2	46	76	0.5	0.5	74	4.5	4	**76.8**	0.8
B3	333	0.9	70	82	0.5	0.5	94	2.5	4	**78.4**	−3.6
R1	176	0.9	40	79	0.5	0.5	51	4	4	**78.8**	−0.2
R2	180	0.5	54	76	0.5	0.5	85	4.5	4	**77.1**	1.1
R3	960	0.8	126	75	0.5	0.5	255	4	4	**74.8**	−0.2
S1	297	0.5	56	67	0.5	0.5	123	9	8	**68.1**	1.1
S2	1176	0.8	106	66	0.5	0.5	288	8	8	**66.3**	0.3

Table 5
Minimum required absorption in relation to signal-to-noise ratio and quality of verbal communication.

Quality of verbal communication	SNR (dB)	$L_{N,A}$ (dB)	$L_{S,A,1m}$ (dB)	Vocal effort	A/N_S (m^2)	A/N (g = 3) (m^2)	A/N (g = 4) (m^2)
Insufficient	−9	83	74	Loud	3.2	1.1	0.8
Insufficient	−6	77	71	Raised	6.3	2.1	1.6
Sufficient	−3	71	68	Raised	12.6	4.2	3.1
Satisfactory	0	65	65	Normal	25	8.4	6.3
Good	3	59	62	Normal	50	17	12.5
Good	6	53	59	Relaxed	100	33	25
Very good	9	47	56	Relaxed	200	67	50

constant power. So, due to the Lombard effect the acoustic treatment of a room is more efficient for the speech noise than for other noise sources.

In addition to the ambient noise level it may be interesting to consider the signal-to-noise ratio. A simple approach is suggested here, namely to define the signal-to-noise ratio as the level difference between the direct sound from a speaking person in a distance of 1 m and the ambient noise in the room. By use of (1) and (8) the signal-to-noise ratio SNR can be expressed in terms of the absorption area per speaking person:

$$SNR = L_{S,A,1m} - L_{N,A} = -14 + 10\log\left(\frac{A}{N_S}\right)(\text{dB}) \quad (9)$$

This applies to a range of the speech levels between 55 dB and 75 dB, or a range of SNR between −10 dB and +10 dB, see Table 5.

6. Guide for the acoustic treatment of eating establishments

For the evaluation of acoustic quality of eating establishments it is suggested to consider the quality of verbal communication which can be related to the signal-to-noise ratio as suggested by Lazarus [8]. Thus SNR between 3 dB and 9 dB is characterized as "good", and the range between 0 dB and 3 dB is called "satisfactory". More examples are shown in Table 5. From (9) it follows that the required absorption area per speaking person can be calculated from:

$$\frac{A}{N_S} = 10^{(SNR+14)/10}(\text{m}^2) \quad (10)$$

In Table 5 the results of such calculations are presented for SNR between −9 dB and + 9 dB in steps of 3 dB. The quality of speech communication is labelled as suggested by Lazarus [8]. The corresponding ambient noise level, speech level, and vocal effort are also presented in Table 5.

From the results in Table 5 it follows that the quality of verbal communication is "insufficient" if $SNR < -3$ dB, or if the absorption area per speaking person is less than 12.6 m^2. With a typical group size g = 3, the minimum absorption area per person is about 4 m^2 for "sufficient" quality of verbal communication. "Satisfactory" verbal communication requires about 8 m^2 absorption area per person, and at least 17 m^2 is required for "good" verbal communication. Assuming a little higher group size g = 4, the minimum absorption area per person is about 3 m^2 for "sufficient" quality of verbal communication, 6.3 m^2 for "satisfactory", and 12.5 m^2 for "good" quality of verbal communication.

The amounts of absorption area required for good quality of verbal communication are obviously very difficult to reach in a room with dense location of tables and seats; in addition to the ceiling area it may be necessary to include some part of the walls and if possible, also the floor for sound absorbing treatment. A high ceiling will also help to achieve the goal if the walls are made sound absorbing.

The results presented in Table 5 can also be used for a room with known absorption area to estimate the maximum number of people in order to keep a certain quality of verbal communication.

7. Conclusion

A new simple prediction model has been derived for the average A-weighted noise level due to many people speaking in a room with assumed diffuse sound field. Thus the prediction model should be used with caution in cases where the sound field cannot be assumed to be diffuse, like large flat rooms with a low ceiling.

It is a precondition that the A-weighted ambient noise level is at least 45 dB. A linear relationship between speech level and ambient noise is assumed. The prediction model has been verified by comparison to measured results published by other researches in two food courts and a canteen, the number of people ranging from ca. 50 to 540. It was found that the Lombard slope should be 0.5 dB/dB in order to achieve good agreement with the measured data. The prediction model has been further tested against results reported from 10 different eating establishments, the number of seats ranging from 40 to 126.

As it is obvious that a gathering of people can be more or less lively and there are more or less unknown social parameters involved, it is necessary to estimate the average group size, defined as the average number of people per speaking person. For the cases of food courts, canteens, bistros, and restaurants the group size parameter was found to be around 3–4, but in a more lively party or a very dead gathering the parameter can take values outside this range; a rather extreme example of this is a group size of 8 suggested for senior residence dining rooms.

Assuming a fixed number of speaking persons in a room, the change of the equivalent absorption area can be expected to lead to a much stronger change of the noise level than for a constant level noise source. Due to the feed-back influence of noise on the speech level (the Lombard effect), the speech level increases in noisy environments, and the suggested prediction model gives a 6 dB reduction of the noise level by doubling the equivalent absorption area of the room. This is in contrast to the 3 dB reduction by doubling of the absorption area for a constant power sound source.

The estimated signal-to-noise ratio and the quality of verbal communication have been used together with the suggested prediction model to establish a design guide for the acoustic treatment of eating establishments. With the assumption of a typical group size of 3, the minimum required absorption area per person is found to be about 4 m^2 for "sufficient" verbal communication, about 8 m^2 for "satisfactory" verbal communication, and as high as 17 m^2 for "good" verbal communication.

References

[1] Lazarus H. Prediction of verbal communication in noise – a review: part 1. Appl Acoust 1986;19:439–64.
[2] ISO 9921. Ergonomics – assessment of speech communication. Geneva; 2003.
[3] Tang SK, Chan DWT, Chan KC. Prediction of sound-pressure level in an occupied enclosure. JASA 1997;101:2990–3.

[4] Kang J. Numerical modelling of the speech intelligibility in dining spaces. Appl Acoust 2002;63:1315–33.
[5] Navarro MPN, Pimentel RL. Speech interference in food courts of shopping centres. Appl Acoust 2007;68:364–75.
[6] Hodgson M, Steiniger G, Razavi Z. Measurement and prediction of speech and noise levels and the Lombard effect in eating establishments. JASA 2007;121:2023–33.
[7] Maekawa Z, Lord P. Environmental and architectural acoustics. London: E & FN Spon; 1994.
[8] Lazarus H. Prediction of verbal communication in noise – a development of generalized SIL curves and the quality of communication (part 2). Appl Acoust 1987;20:245–61.