

SPEECH BABBLE AS SOUND MASKING IN OPEN-PLAN OFFICES – A STATISTICAL APPROACH

P Leonard Max Fordham LLP
A Chilton Max Fordham LLP

1 INTRODUCTION

A study has been undertaken to establish the conditions under which speech babble can provide useful masking noise in open-plan office settings.

Previous studies have shown that the distraction caused by speech diminishes when there are multiple simultaneous voices – effectively creating an unintelligible ‘babble’. Where speech babble is not distracting to work, it can potentially provide useful masking sound, improving privacy. This is more likely to be the case for higher occupancy offices and for work activities involving a greater proportion of speech.

A simple statistical model based on a binomial distribution of talkers has been used to determine the minimum office occupancy for which speech babble masking may be effective. The same model is used to determine the quantity of acoustic absorption per occupant that would be consistent with effective speech babble masking.

2 BASIS OF APPROACH

2.1 Motivation

The motivation for this study was the authors’ personal experience of working in the various Max Fordham open-plan offices around the country. The subjective impression is that in the larger (higher occupancy) main office, the open-plan work environment provides acceptable speech privacy and distraction conditions despite the absence of masking sound from either mechanical services or electronic masking systems. However, in the smaller (lower occupancy) offices, this is not the case.

The hypothesis is that the noise generated by the occupants can provide an appropriate level of masking sound, but this is only likely to be successful if certain conditions are met. This study seeks to investigate the key parameters involved in determining whether successful conditions are met and to provide a simple framework that could be used for design of open-plan offices.

2.2 Speech Babble as Masking Sound

A study by Jones & Macken [1] on Auditory Babble and Cognitive Efficiency investigated the disruptive effect of irrelevant speech babble on short-term memory. They found that the degree of disruption due to speech babble started to decrease (relative to that caused by a single voice) when the number of simultaneous voices increased beyond three. They speculate that this is because it might still be possible to ‘perceptually separate or stream out’ the individual voices when the auditory stimulus comprised a small number of voices but this is not possible as the number of voices increases. They conclude that *“a certain degree of clamor, say, in a large office, will be less disruptive than in a smaller office, in which few sources of sound do not mask each other”*.

A review of research on speech intelligibility in multiple-talker conditions [2] showed that speech is not an effective masker and to achieve a 50% intelligibility level, interfering speech needs to be almost 10dB louder than steady-state noise. It was noted that this release from masking disappears when the number of interfering talkers is increased and that this occurs in a number of studies when there are four or more talkers.

Therefore, we can say that if there are a sufficient number of interfering voices then they will no longer be distracting and will indeed perform the function of sound masking. The required number is likely to be at least 4 but may depend upon the spatial location of the voices.

2.3 Masking Sound Level

If speech babble is to act as a masking sound within an office environment, then it must be at a high enough level such that it can effectively mask a speech signal.

BS EN ISO 3382-3:2012 [3] provides a model for the effect of speech intelligibility on the performance of cognitively demanding tasks and states that “the negative effects of speech on work performance start to vanish rapidly if the STI is below 0.5”.

STI is difficult to predict with accuracy, however signal to noise ratio can be used as a proxy for STI and is a much simpler parameter to work with in a basic model.

A study by Lazarus [4] provides a correlation of signal to noise ratio with STI and the value of which are reproduced in Table 1.

Table 1: Correlation of Speech Intelligibility with Signal to Noise Ratio, taken from Lazarus [4]

STI	0.1	0.2	0.3	0.4	0.5	0.6
SNR, dB	-12	-9	-6	-3	0	3

In a study of the Speech Transmission Index in large spaces Liu, Maa, Kang & Wanga [5] measured the relation between speech intelligibility scores and the signal to noise ratio. The correlation coefficient R^2 between the two parameters was 0.927 and speech intelligibility crossed the threshold of 50% at a signal to noise ratio of ~0dB.

It is suggested that distraction due to direct speech sounds is unlikely if the level of direct speech rarely exceeds the ambient noise level.

2.4 Noise Levels

High ambient noise levels need to be avoided in open-plan offices to prevent discomfort, interference with tasks requiring concentration and ultimately speech communication between occupants.

The French Standards Institute [6] recommend different average levels of ambient noise depending on the type of open-plan office. These are summarised in Table 2.

Table 2: Target average activity noise levels from NFS 31-199 [6]

Activity Type	Average ambient noise level, $L_{Aeq,T}$
Telephone work	48-52
Collaborative work	45-50
Partially collaborative work	40-45

The values in Table 2 relate to average noise levels. It is reasonable to expect that noise levels can be higher in the short term without causing excessive interference to working tasks but the authors are not aware of any specific information covering this situation.

Reference [7] gives an overview of studies measuring ambient noise level in open-plan offices and contact centres. The measurements are summarised in Table 3.

As an upper limit, it is suggested that short-term noise levels should not be so high as to have an unacceptable effect on speech communication between nearby co-workers. Table 4 shows the maximum steady noise levels for reliable speech communication at a distance of 1m, taken from British Standard 8233 [8]. For the purposes of this study, a level of 62dBA is adopted as the level above which the effect on speech communication is likely to be unacceptable.

Table 3: Overview of studies measuring ambient noise in open-plan offices and contact centres [7]

Type of Office	No. Offices	Measured Ambient Noise Level (dBA)	
		Range	Average
Typical Open-Plan Office	20	41-53	47
Contact Centre	21	50-62	58

Table 4: Maximum steady noise levels for reliable speech communication at 1m (from BS 8233 [8])

Voice Level	Maximum steady noise level, $L_{Aeq,T}$
Normal voice	57
Raised voice	62

2.5 Conditions to be Satisfied

It is suggested that the following conditions are necessary for speech babble to provide effective masking:

Condition 1: It must be the case for most of the time that either a minimum of four people are talking or alternatively that no-one is talking. This is based on the finding¹ that at least three voices must be present simultaneously to act as an unintelligible masking babble. A minimum of four talkers are required such that at least three voices are present to mask the fourth.

Condition 2: The ambient noise level should be less than 62dBA for most of the time. This condition is intended to avoid unacceptable interference with speech communication at 1m. It may be that a lower noise limit may be considered appropriate, in which case the calculations can be adjusted accordingly.

Condition 3: It should be the case for most of the time that there is no one talking within a distance of a particular listener that would result in the direct speech sound level exceeding the total reverberant sound level at the listener. This is to avoid excessive speech disturbance and is similar to the concept of distraction distance [3]. Again, it would be possible to adapt this approach to a different signal-to-noise threshold.

It is important to note that each condition introduces a threshold. These relate respectively to the minimum number of simultaneous talkers, the ambient noise level and the limiting signal-to-noise ratio for distraction. These thresholds are put forward by the authors of this paper for the purposes of developing the analytical equations into design relationships. It would be equally possible to develop the approach with a different set of thresholds.

It is also important to note that each condition is to be met 'most of the time' as opposed to 'always' or 'on average'. This relates to the statistical nature of open-plan office acoustics which are dependent on a population of occupants. This statistical basis is a central aspect of the approach presented in this paper. The threshold for 'most of the time' is again put forward by the authors (as detailed in the subsequent sections). The methodology can again be adapted to alternative thresholds if considered more appropriate.

2.6 Simplifying Assumptions

The main simplifying assumptions of the model presented are as follows:

- 1) It is assumed that all occupants of the office are identical in terms of their speech level and the amount of time spent talking. This is clearly a simplification of a real environment where there will be significant diversity. The assumption is made to allow a simple statistical model to be used.
- 2) Sources of noise other than speech are not included in the model. It was noticeable in studies of our own office environment that noise from keyboard typing and computer equipment are significant.
- 3) For the purposes of evaluating ambient noise level, the office is treated as a reverberant space. This assumption is made because it significantly simplifies the calculations, whilst acknowledging that large office spaces with highly absorbent ceilings are not fully reverberant spaces.
- 4) It is assumed that the Lombard Effect is applicable i.e. occupants increase their speech level in response to higher ambient noise levels. This was found to be the case in a previous study [7] by the same authors.
- 5) Occupants are randomly distributed throughout the space.

3 CONDITION 1 – MINIMUM NUMBER OF SIMULTANEOUS TALKERS

3.1 Evaluating Minimum Occupancy

In order to estimate the probability that at least four people are talking simultaneously, we can assume that each of the office occupants speaks, on average, for a fraction, t , of the time. The value t is termed “average fractional talk time”. For example, an occupant who speaks for 12% of the time on average during the course of their work has $t = 0.12$.

The total occupancy of the office is N .

The probability, $P(N_s = k)$, of the number of simultaneous talkers, N_s , being exactly k at any given time can be evaluated using a binomial distribution function.

$$P(N_s = k) = \left(\frac{N!}{k!(N-k)!} \right) \cdot t^k \cdot (1-t)^{N-k} \quad \text{Eq. 1}$$

For Condition 1 to be met, the number of simultaneous talkers must be either zero or at least four. We are therefore looking to evaluate the cumulative total probability, P_{babble} , that k is any value excluding 1, 2 and 3.

$$P_{babble} = 1 - P(N_s = 1) - P(N_s = 2) - P(N_s = 3) \quad \text{Eq. 2}$$

Where: $P(N_s = 1)$, $P(N_s = 2)$, and $P(N_s = 3)$ are evaluated using Eq. 1.

We now require that this condition should be met for ‘most of the time’ so that speech babble can generally be relied on to provide masking. We can do this by evaluating the conditions for which P_{babble} exceeds 0.9 (i.e. 90% of the time).

Given that P_{babble} is a function only of N and t , we can evaluate the minimum office occupancy, N_{min} , for which P_{babble} will exceed 0.9. These values are tabulated in Table 5 for various values of t .

Table 5: Minimum office occupancy, N_{min} , for which speech babble masking can be effective, shown for various values of fractional talk time, t

t	0.050	0.075	0.100	0.125	0.150	0.200	0.300	0.400
N_{min}	132	87	65	52	43	32	21	15

It can be seen that speech babble cannot be relied upon to provide speech masking unless there are a sufficient number of people in the office. Unsurprisingly, the minimum occupancy required decreases as the amount that people tend to speak increases.

3.2 Developing an Analytical Solution

The tabulated values in Table 5 were calculated using Eq. 1 and Eq. 2 but they do not give a direct analytical relationship between N_{min} and t . In order to find an approximate analytical solution, an alternative definition of ‘most of the time’ can be used. In this method, it is proposed that the mean of the probability distribution minus one standard deviation should be equal to four. With reference to Figure 1, there is a large proportion of the area under the probability graph (i.e. the cumulative probability) associated with the number of simultaneous talkers being at least four. Note that this method ignores the situation where there are no speakers at all ($k = 0$).

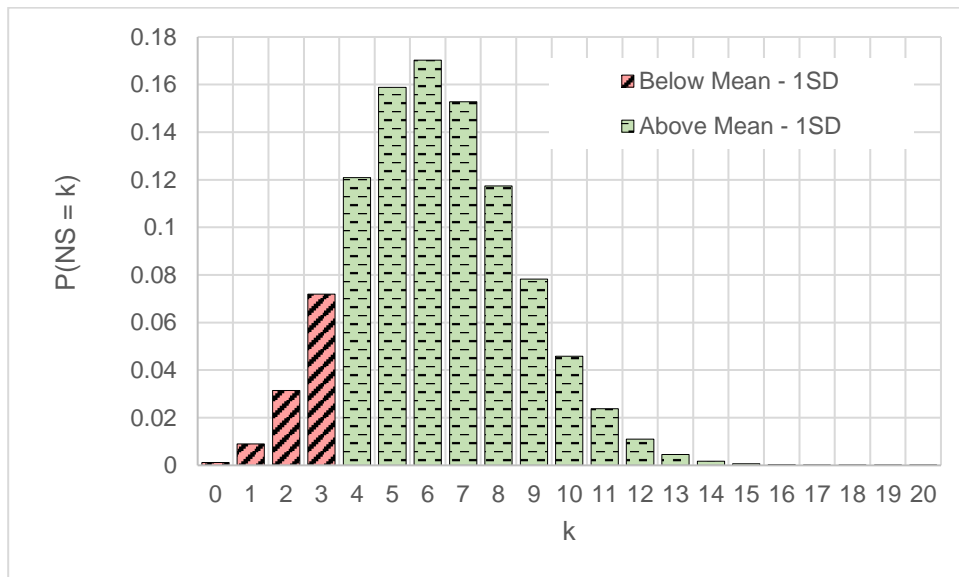


Figure 1: Binomial probability distribution for $N=50$ and $t=0.125$ showing the proportions of the distribution lying either side of one standard deviation (1SD) below the mean. In the example distribution shown, 1SD below the mean occurs at a value of around 3.91

For the binomial distribution described by Eq. 1, the mean is equal to $N \cdot t$ and the variance is equal to $N \cdot t \cdot (1 - t)$. Noting that the standard deviation is the square root of the variance, we get the following equation.

$$N_{min} \cdot t - \sqrt{N_{min} \cdot t \cdot (1 - t)} = 4 \quad \text{Eq. 3}$$

This can be rearranged to a quadratic format and solved as:

$$N_{min} = \frac{9 - t + \sqrt{(1 - t) \cdot \sqrt{(17 - t)}}}{2t} \quad \text{Eq. 4}$$

Given that t is small (unlikely to exceed 0.5), $\sqrt{(1-t)}$ can be approximated by the first two terms of its Taylor Series $(1 - \frac{t}{2})$. Also using $(17-t) \approx 17$, we can produce the following approximated version of Eq. 4.

$$N_{min} \cong \frac{6.56}{t} - 1.53 \quad \text{Eq. 5}$$

Comparing the values given by Eq. 5 to those evaluated previously using Eq. 1 and Eq. 2 shows that this is a useful, simplified approximation.

Table 6: A comparison between $N_{min,approx}$ values calculated using Eq. 5 and those, N_{min} , calculated long-hand using Eq. 1 and Eq. 2

t	0.050	0.075	0.100	0.125	0.150	0.200	0.300	0.400
N_{min} (using Eq. 1 & Eq. 2)	132	87	65	52	43	32	21	15
$N_{min,approx}$ (using Eq. 5)	130	86	64	51	42	31	20	15

4 CONDITION 2 – LIMIT ON AMBIENT NOISE LEVEL

Using the model developed by Rindel [9], and assuming a Lombard coefficient (slope) of 0.5, the ambient noise level, $L_{N,A}$, can be evaluated using the following relationship:

$$L_{N,A} = 93 - 20 \log\left(\frac{A}{N_S}\right) \quad \text{Eq. 6}$$

Where N_S is the number of people speaking simultaneously.

To evaluate an ambient noise level that is not normally exceeded, $L_{N,A[M+1SD]}$, we will use a value for N_S that is one standard deviation above the mean number of simultaneous talkers. With reference to the discussion leading to Eq. 3, we get:

$$N_{S[M+1SD]} = N \cdot t + \sqrt{N \cdot t \cdot (1-t)} \quad \text{Eq. 7}$$

Where $N_{S[M+1SD]}$ is the number of simultaneous speakers a one standard deviation above the mean, N is the total occupancy of the office and t is the fractional talk time.

Rearranging Eq. 6 and using Eq. 7, we get:

$$A = \left[N \cdot t + \sqrt{N \cdot t \cdot (1-t)} \right] \cdot 10^{\left(\frac{93 - L_{N,A[M+1SD]}}{20} \right)} \quad \text{Eq. 8}$$

or, alternatively:

$$\frac{A}{N} = \left[t + \sqrt{\frac{t \cdot (1-t)}{N}} \right] \cdot 10^{\left(\frac{93 - L_{N,A[M+1SD]}}{20} \right)} \quad \text{Eq. 9}$$

Condition 2 is that the ambient noise level does not normally exceed 62dBA. This gives us the following expression for the minimum quantity of absorption (expressed per occupant):

$$\frac{A_{min}}{N} \approx 35.5 \left[t + \sqrt{\frac{t \cdot (1-t)}{N}} \right] \quad \text{Eq. 10}$$

Table 7: Minimum Sabines per occupant required to keep ambient noise levels within 62dBA for most of the time. Values shown for different values of t and at extreme values of N .

t	0.050	0.075	0.100	0.125	0.150	0.200	0.300	0.400
A_{\min}/N (for $N=N_{\min}$)	2.4	3.7	4.9	6.1	7.3	9.6	14.2	18.7
A_{\min}/N (for $N \rightarrow \infty$)	1.8	2.7	3.5	4.4	5.3	7.1	10.6	14.2

5 CONDITION 3 – LIMIT ON SPEECH DISTRACTION

Using the model developed by Rindel [9], and assuming a Lombard coefficient (slope) of 0.5, the ambient noise level, $L_{N,A}$ can be described by Eq. 6.

The estimated speech level at distance r , considering the Lombard effect, is described by Eq. 11.

$$L_{S,A,r} = 55 + c(L_{N,A} - 45) - 20 \log r \quad \text{Eq. 11}$$

For any given number of simultaneous speakers N_S , where $N_S > 0$, the ambient noise level from people speaking is given by Eq. 6 in the previous section.

The distance, $r_{SNR=0}$, at which the direct speech level is equal to the reverberant speech level (Signal to Noise, SNR, equal to 0dB) can be calculated by setting Eq. 6 equal to Eq. 11 and making r the subject.

$$(r_{SNR=0})^2 = \frac{A \cdot 10^{-1.4}}{N_S} \quad \text{Eq. 12}$$

The expected number of people talking, $E(N_{SNR \geq 0})$, within the area given by radius $r_{SNR=0}$ is found by multiplying that area by the number of people speaking per unit area ($N_S / N \cdot S_{pp}$) where S_{pp} is floor area per person, or occupancy density.

$$E(N_{SNR \geq 0}) = \frac{\pi \cdot (r_{SNR=0})^2 \cdot N_S}{N \cdot S_{pp}} \quad \text{Eq. 13}$$

Substituting Eq. 12 into Eq. 13 and simplifying ($1/\pi \cdot 10^{-1.4} \cong 8$).

$$E(N_{SNR \geq 0}) = \frac{A}{8 \cdot S_{pp} \cdot N} \quad \text{Eq. 14}$$

Assuming a Poisson distribution, the probability that the number of people talking within the radius is equal to zero is:

$$P(N_{SNR \geq 0} = 0) = e^{-E(N_{SNR \geq 0})} \quad \text{Eq. 15}$$

This accounts for the case where the number of people speaking within the room, N_S , is >0 , but the number of people speaking within the radius, $r_{SNR=0}$, is 0. The case for the number of people speaking within the room, N_S , being zero (where, by definition, the number of people speaking within the radius is also zero) must also be accounted for.

With reference to Eq. 1 the probability of there being no talkers within the room is:

$$P(N_S = 0) = (1 - t)^N \quad \text{Eq. 16}$$

Therefore, the total probability of there being no talkers within the radius, including the case of there being no talkers within the room as whole is equal to:

$$P(N_S = 0) + (1 - P(N_S = 0)) \cdot P(N_{SNR \geq 0} = 0) = (1 - t)^N + (1 - (1 - t)^N) \cdot e^{-E(N_{SNR \geq 0})} \quad \text{Eq. 17}$$

For a given value of N , t and S_{pp} , the maximum value of A/N can be found such that the probability of no-one speaking within the radius $r_{SNR=0}$ achieves some specific value. An initial suggestion is to achieve this for 90% of the time. Setting Eq. 17 to equal 90% and rearranging for A_{max}/N :

$$A_{max}/N = -S_{pp}(8\ln(0.9 - (1 - t)^N) - \ln(1 - (1 - t)^N)) \quad \text{Eq. 18}$$

Table 8: Maximum Sabines per occupant required to limit speech distraction within direct sound distraction radius for most of the time. Values shown for different values of t and at extreme values of N . Calculated for $S_{pp}=10\text{m}^2$

t	0.050	0.075	0.100	0.125	0.150	0.200	0.300	0.400
$S_{pp}=10\text{m}^2$ A_{max}/N (for $N=N_{min}$)	8.5	8.5	8.5	8.5	8.5	8.5	8.5	8.5
A_{max}/N (for $N \rightarrow \infty$)	8.4	8.4	8.4	8.4	8.4	8.4	8.4	8.4

The value for A_{max}/N does not meaningfully change for varying values of t or N and the following simplifications can be made:

Given that t is small (unlikely to exceed 0.5) and N is large (usually greater than 10) then $(1 - t)^N$ is negligible. Also noting that $\ln(1) = 0$ Eq. 18 simplifies to:

$$A_{max}/N = 0.84 \cdot S_{pp} \quad \text{Eq. 19}$$

Therefore, in order to ensure that $A_{max} > A_{min}$ then a requirement for S_{pp} can be defined as:

$$S_{pp_{min}} = 42.3 \left[t + \sqrt{\frac{t \cdot (1 - t)}{N}} \right] \quad \text{Eq. 20}$$

6 CONCLUSIONS

By using a simple statistical model and setting out three conditions under which speech babble may provide effective masking in an open-plan office, it has been possible to establish design limits on the minimum occupancy, N_{min} , the minimum floor area per person $S_{pp_{min}}$ and on the amount of acoustic absorption per occupant, A/N . These limits could potentially be used to inform the design of open-plan offices.

For example, if providing a new open-plan office space for a known number (and talk time behaviour) of occupants, the sequential design steps would be as follows:

1. Check if N exceeds N_{min} (using Eq. 4). If it does, then continue to Step 2. If it does not, then provide a quantity of acoustic treatment exceeding A_{min}/N and consider the use of electronic or mechanical services sound masking.
2. Check if S_{pp} exceeds $S_{pp_{min}}$ (using Eq. 20). If it does, then continue to Step 3. If it does not, then either increase the office floor area appropriately and continue to Step 3 or provide a quantity of acoustic treatment exceeding A_{min}/N and consider the use of electronic or mechanical services sound masking. (See also Note A, below).
3. Provide a quantity of acoustic treatment exceeding A_{min}/N but not exceeding A_{max}/N (refer to Eq. 10 and Eq. 19). (See also Note A, below).

Note A: This study takes no account of the effect of barriers or screening, which may offer alternative design solutions. For example, with appropriate use of barriers, it may not be necessary for S_{pp} to exceed $S_{pp_{min}}$ and may be acceptable for A_{max}/N to be exceeded.

It is suggested that further work should be undertaken to validate the model against real life situations, preferably using post-occupancy evaluation to determine if occupants consider the acoustics to adversely affect their work activities. Validations of this type will inform the thresholds applied to each of the three conditions on which the design relationships are based.

7 REFERENCES

- [1] D. M. Jones and W. J. Macken, "Auditory Babble and Cognitive Efficiency: Role of Number of Voices and Their Location," *Journal of Experimental Psychology Applied*, vol. 1, no. 3, pp. 216-226, 1995.
- [2] A. W. Bronkhorst, "The Cocktail Party Phenomenon: A Review of Research on Speech Intelligibility in Multiple-Talker Conditions," *Acta Acustica*, vol. 86, pp. 117-128, 2000.
- [3] BSI, *BS EN ISO 3382-3:2012: Acoustics - Measurement of Room Acoustics Parameters*, 2012.
- [4] H. Lazarus, "Prediction of verbal communication in noise - A development of generalized SIL curves and the quality of communication (Part 2)," *Applied Acoustics*, vol. 20, no. 4, pp. 245-261, 1987.
- [5] H. Liu, H. Ma, J. Kang and C. Wang, "The speech intelligibility and applicability of the speech transmission index in large spaces," *Applied Acoustics*, vol. 167, 2020.
- [6] AFNOR, *NF S31-199: 2016. Acoustique – Performances acoustiques des espaces ouverts de bureaux.*, 2016.
- [7] P. Leonard and A. Chilton, "The Lombard Effect in Open Plan Offices," in *Proceedings of the Institute of Acoustics*, 2019.
- [8] BSI, *BS 8233:2014: Guidance on sound insulation and noise reduction for buildings*, 2014.
- [9] J. H. Rindel, "Verbal communication and noise in eating establishments," *Applied Acoustics*, vol. 71, no. 12, pp. 1156-1161, 2010.