Proceedings of The Institute of Acoustics

COMPARISON OF THE STEADY STATE AND DYNAMIC CHARACTERISTICS OF SOUND INSULATION WITH THE INTELLIGIBILITY OF TRANSMITTED SPEECH

P.T. LEWIS

Welsh School of Architecture, UWIST, Cardiff.

INTRODUCTION

The question of the intelligibility of speech in buildings is treated differently depending on whether the speech is spoken and heard within the same room, e.g. an auditorium, or has been transmitted from one room to another. In auditoria, it is considered to be a problem in the time domain with the emphasis on the rate of sound decay expressed in overall terms by the reverberation time and more precisely by the proportion of the energy that is received within the integration time of the ear. In contrast, the intelligibility of speech when transmitted between rooms is treated as a frequency domain problem and is taken to be due to the difference in mean sound pressure levels in source and receiving rooms expressed as a function of frequency.

However, if we compare the sound decay patterns found in some sound transmission situations with those of auditoria whose reverberation times are generally considered to be much too long for good speech intelligibility (Figure 1), it is apparent that there can be as much if not more 'late' energy in the former case as in the latter. The aim of the work described in this paper was to assess how significant this phenomenon could be and, in particularly, whether walls and partitions with the same steady state sound insulation but with different dynamic characteristics could produce different levels of speech privacy.

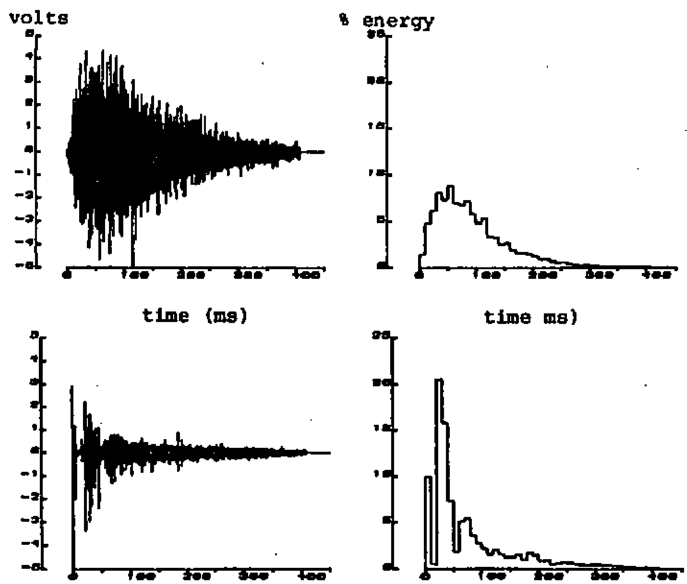
FIGURE 1

Impulse response of sound transmission between rooms

78% late energy

Impulse response of concert hall (RT = 2.2s)

46% late energy



COMPARISON OF THE STEADY STATE AND DYNAMIC CHARACTERISTICS OF SOUND INSULATION WITH THE INTELLIGIBILITY OF TRANSMITTED SPEECH

COMPUTER SIMULATION OF TRANSMISSION

Signal Processing

In order to achieve the degree of control over the signals that was desirable for this study, the transmission process was simulated using a digital signal processing technique. This procedure, which is described in more detail in Reference [1] consists essentially of

(a) passing an analogue recording of anechoic speech through an analogue to digital converter to produce a digital file

(b) repeating this process for the impulse response of the transmission path

(c) 'convolving' the two signals to produce a digital file of the output signal, i.e. the signal that would be received if the anechoic speech had been transmitted along the respective path, using the well known integral

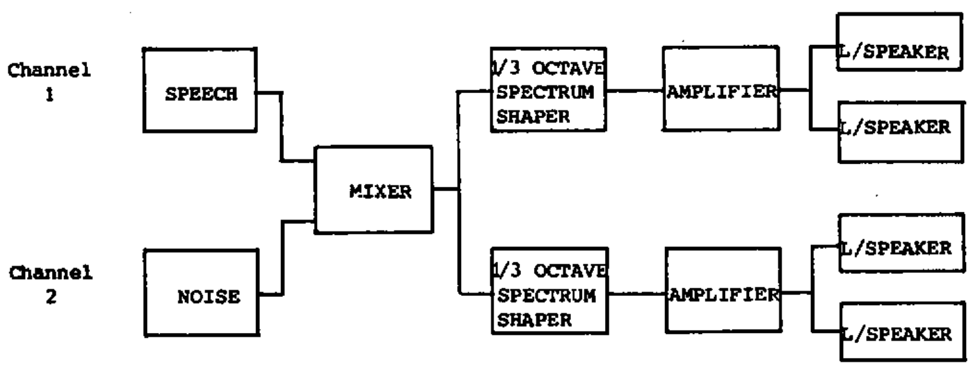
$$y(t) = \int_{-\infty}^{t} x(t)\ h(t-\tau)\ d\tau \qquad (1)$$

(d) converting the digital output file to analogue form by passing it through a digital to analogue converter.

Simulation of sound field

Ideally, in order to realistically simulate the sound field that would occur in field situation, the speech signals would need to be replayed so that the amplitude, time pattern, frequency composition and direction of energy arrival of the sound were faithfully represented. Unfortunately, the processing was limited to a single channel, loudspeakers are not perfect and the room available although very 'dead' was not fully anechoic so that some compromises were inevitable. Nonetheless, using the equipment shown in Figure 2, many of these deficiencies were substantially reduced if not eliminated and the listening conditions created were more than adequate for this work.

FIGURE 2 block diagram of equipment used for speech intelligibility tests



COMPARISON OF THE STEADY STATE AND DYNAMIC CHARACTERISTICS OF SOUND INSULATION WITH THE INTELLIGIBILITY OF TRANSMITTED SPEECH

In order to help create a sense of hearing the speech in a normal room and also to correct for the deficiencies in the frequency response of the loudspeakers and for room modes, four loudspeakers were used situated near the corners of the listening room. The two front loudspeakers were operated at a higher amplitude to create the impression that the sound radiated from their direction and were fed via a 1/3 octave spectrum shaper to flatten the peaks in their own and the rooms' frequency responses (principally at low frequencies). The rear speakers, in conjunction with their 1/3 octave spectrum shaper, were used to give a flat overall spectrum by filling in the troughs in the frequency response obtained with the front speakers and to create the impression of being in a sound field rather than listening to loudspeakers.

SPEECH INTELLIGIBILITY TESTS

Selection of test

The range of tests suitable for this study was limited by the fact that it was not the intelligibility of individual words in isolation that was of interest but rather how the clarity of these words was affected by residual/late energy from previous words or speech sounds. The need therefore was for the test words or sounds to be embedded in a sentence or for the syllable/sound of interest to occur at the end of a word in order to allow the masking process caused by late energy to take effect.

These requirements were met by the FAAF2 test developed by Forster and Haggard. [2] Essentially it consists of 80 single syllable words divided into 20 sets of four. The words within each set differ only in their initial or final consonant sounds. For example:

mail	bail	nail	dale
bang	bad	bag	ban

One of the words in each group is presented to the subject in the carrier sentence

Can you hear (X) clearly?

and the task is to identify (X) from the list of four alternatives.

A complete 80 item test consists of four response pages, each page containing the full list of 80 words but with a varying order of the four word sets coupled with a random position of the target word.

Test recordings.

Measurements of the impulse responses of a range of transmission situations were obtained and, from the results, three were selected for the subjective tests - two of which exhibited a high proportion of late energy and one with a low proportion. Each impulse response was then convolved with an 80 word list to produce a test recording for presentation to a panel of subjects. The anechoic version of the list was also included as a reference giving a total of four levels of late energy.

Speech intelligibility tests

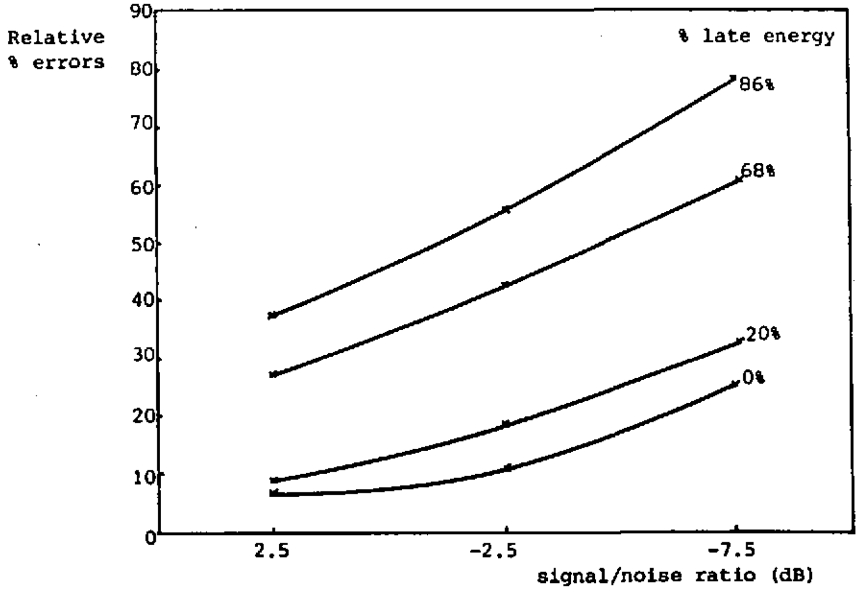
The FAAF2 test is designed to be carried out in the presence of speech-spectrum shaped noise and, by varying the relative levels of the speech and the background noise, the average number of words successfully recognised in a test is changed accordingly. This feature allowed the trade-off between fraction and signal to noise ratio to be examined.

A sample of 15 subjects each carried out an 80 word test for each of the four different values of LEF in three different levels of background noise, making a total of 12 tests. The speech level was chosen to approximately that which would be heard from conversation carried out at normal level and transmitted through a wall with a steady state sound insulation of 35 dB i.e. an rms speech level of 32.5dB. The background noise levels were chosen to give a reasonable spread of results in the FAAF test and corresponded to (speech - background) differences of +2.5dB, -2.5dB and -7.5dB, i.e. the noise levels used were 32.5dB, 37.5dB and 42.5dB respectively.

RESULTS

Figure 3 shows the results that were obtained in terms of the relative percentage of errors plotted against the speech/noise ratio for the four values of late energy fraction.

Figure 3



The nature of the FAAF test is such that, even if no words were recognised, an average of 20 correct answers out of a possible 80 would be obtained due to chance. The ordinare in Figure 3 is therefore calculated from

$$\text{relative \% errors} = \frac{\text{actual errors}}{60} \times 1000 \qquad (2)$$

COMPARISON OF THE STEADY STATE AND DYNAMIC CHARACTERISTICS OF SOUND INSULATION WITH THE INTELLIGIBILITY OF TRANSMITTED SPEECH

If the time pattern of energy arrival had no effect on speech intelligibility, all curves in Figure 3 would be identical to the results for the anechoic signal. However, it is clear that this is not so and that the more energy that arrives late, i.e. >50msec after the first energy received, the more intelligibility is reduced.

IMPLICATIONS

These results suggest that steady state sound insulation is only a partial indicator of the level of speech privacy that will be experienced. Situations with the same steady state sound insulation and the same level of background noise can produce significantly different levels of speech intelligibility if they differ in the proportion of energy which arrives later than the integration time of the ear for speech.

It would appear that a change in the late energy fraction of approximately 60% has the same effect on the intelligibility of transmitted speech as a change of 10dB in the steady state sound insulation.

More work is required to identify those parameters on which the late energy fraction depends.

REFERENCES

[1] P.T. Lewis and C.B. Chinoy 'Development of a computer-based building acoustics simulator', Proceedings of the IOA Spring Conference 1984, Swansea.

[2] J.R. Forster and M.P. Haggard 'Introduction and test manual for FAAF II : the four alternative auditory feature test (Mark 2)' MRC Institute of Hearing Research, University of Nottingham.